

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 06:01:15 ; Search time 5504.92 Seconds  
(without alignments)  
8515.190 Million cell updates/sec

Title: US-09-515-369b-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatacgaactactatagg...aatctccacaagccttgc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sta:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Length	ID	Description
------------	--------------------	----	-------------

	1	2209	98.6	2256	9	AF217405	Homo sapi
	2	2162.6	96.5	193317	2	AC098935	Homo sapi
c	3	2162.6	96.5	229302	2	AC023534	Homo sapi
	4	2105.4	94.0	180048	2	AC068122	Homo sapi
	5	1826.2	81.5	11114	9	AF062931	Homo sapi
	6	1319.2	58.9	7025	9	AF276916	Homo sapi
	7	91.8	4.1	29034	2	HS243947	Homo sapi
c	8	91.8	4.1	174210	2	AC108684	Homo sapi
c	9	91.2	4.1	137769	9	AC002090	Genomic S
	10	90.8	4.1	193126	9	AL139275	Human DNA
	11	90	4.0	177957	2	AC025298	Homo sapi
c	12	90	4.0	178477	2	AC022801	Homo sapi
	13	90	4.0	192695	2	AC098645	Homo sapi
	14	89.2	4.0	181009	2	AC011060	Papio cyn
	15	89.2	4.0	183483	9	AL161628	Homo sapi
	16	89.2	4.0	194590	2	HS315114	Human DNA
	17	88.4	3.9	168486	9	CNS057BN	Human chr
	18	88.2	3.9	149973	2	AC023392	Homo sapi
	19	88.2	3.9	158905	9	AC005666	Homo sapi
c	20	88.2	3.9	163242	2	AC068438	Homo sapi
c	21	87.4	3.9	181134	2	AC074343	Homo sapi
c	22	87	3.9	62497	2	AC103865	Homo sapi
c	23	87	3.9	178756	2	AC104420	Homo sapi
c	24	86	3.8	116516	2	AC109809	Homo sapi
c	25	85.4	3.8	86025	9	AL158175	Human DNA
c	26	85.4	3.8	119361	2	AC092771	Homo sapi
c	27	85.4	3.8	142192	2	AC092771	Homo sapi
c	28	85.4	3.8	158341	2	AC021733	Homo sapi
29	85.4	3.8	160671	9	AC090451	Homo sapi	
c	30	85.4	3.8	162625	9	AC006083	Homo sapi
c	31	85	3.8	127936	2	AC022099	Homo sapi
c	32	85	3.8	143536	2	AC020846	Mus muscu
c	33	85	3.8	172830	9	AL137072	Human DNA
c	34	84.6	3.8	59319	2	AC103842	Homo sapi
c	35	84.6	3.8	114467	9	AP000462	Homo sapi
c	36	84.6	3.8	154236	2	AC051650	Homo sapi
37	84.6	3.8	157322	9	AC012103	Homo sapi	
c	38	84.6	3.8	157399	2	AP003167	Homo sapi
c	39	84.6	3.8	170965	9	AB017654	Homo sapi
c	40	84.4	3.8	144714	9	AC020891	Homo sapi
41	84.4	3.8	145886	9	AP003438	Homo sapi	
42	84.4	3.8	175830	2	AC090547	Homo sapi	
c	43	84.4	3.8	188764	2	AC021238	Homo sapi
c	44	84	3.8	170893	2	AC027537	Homo sapi
c	45	83.8	3.7	37321	9	AC005498	Homo sapi

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	MEDLINE	JOHNNAL
1	AF217405	2256 bp DNA linear PRI 19-SEP-2000	AF217405				human.	1	Madireddi, M.T., Dent, P. and Fisher, P.B.	promoter activity during human melanoma differentiation	J. Cell. Physiol. 185 (1), 36-46 (2000)	10942517		JOHNNAL
	AF217405	2256 bp DNA linear PRI 19-SEP-2000	AF217405				human.	2	Madireddi, M.T. and Fisher, P.B.	Direct Submission	Submitted (17-DEC-1999) Urology, Columbia University, 630 West,			JOHNNAL

168th Street, BB 1505, New York, NY 10032, USA

FEATURES  
source  
1..2256  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tisue\_type="Placenta"  
promoter  
1..2208  
/gene="ST16"  
1..>2256  
/gene="ST16"  
/note="mda-7"  
2209..>2256  
/gene="ST16"  
/product="suppression of tumorigenicity 16 protein"  
mRNA  
BASE COUNT 539 a 584 c 503 g 630 t  
ORIGIN

Query Match 98.6%; Score 2209; DB 9; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 tcaaccttggaaaccagagctgcctgcctcccaaaagctgtactcaataactatattca 91  
DB 1 TCACCTTTTGAACCCAGGCTCTGCTGCTCCCAAAAGCTGTACTCATTAAGATTCTCAA 60

QY 92 ctgatacttgagccaagagttccctagttctcctcctgaacctccctcctgaagtaataagc 151  
DB 61 CTGATTTTGGGCCAAGCTTCCTAGGTTCTCTGCTTGAACCTCTTCTGTAAGTAATATGC 120

QY 152 tatgataagctcaatcggaagctgaagccagagcacaagtcttgccctgaactatcagtta 211  
DB 121 TATGATTAAGCTCATCGAGAGCTGAGGCCAGGACATGTTTGCTGTAACATTCATGTTA 180

QY 212 tatgattcctctctcagagaagatgagctacacagatcccaaggtgtaccctgaagccag 271  
DB 181 TATGATTCTCTCTCCTCAGACAGAGTGAAGTACTACAGATCCAGGISTACCTTGAGGCCAG 240

QY 272 ccaaggctgatacctcagctcagctcctcctcctcagcctcctcctcctcctcctcctcctc 331  
DB 241 CCAAGGTGATATCCATGACGTCATGCTCTGTTTCCAGCTGACCTGTTTAAACAGCTCATGCCA 300

QY 332 cctgcctgcctcctccgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 391  
DB 301 CCTGCTGCTCCTCCCGCTATCTGACAGACAGTATGAGGATTTAGCTGCTGCGG 360

QY 392 ctaattcctcctcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 451  
DB 361 CTGATTTTCCCTCTCAGCTTCCTGCTTACGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 452 ccagacatcctcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 511  
DB 421 CCAGCACTCTCAGCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 512 atagtagttagttagaggtctcctgaatgctcctcctcctcctcctcctcctcctcctcctc 571  
DB 481 ATGAGTATTAAGTTAGGGGCTTCTGTAATGCTTAAATCCACATGCTGGAAGGGGGAG 540

QY 572 tgggggaagagagctgcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 631  
DB 541 TGGGGCAAGAGAGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 632 ccaggaacaaagagctcagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 691  
DB 601 CCAGGAACAAAGGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 692 tgcgaagagagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 751  
DB 661 TGCAGAAAGGATTTCTTGTGCTTAAAGAAATTAACAACTGATGATGAGATGAGGAGGA 720

QY 752 ggggggtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 811  
DB 721 GGGGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 812 cctcttcttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 871  
DB 781 CCTCTTTTCTTTCATCTCCTCAGCTGAGCGAGCTTCTGTCCTCCCGACAGCTACATTC 840

QY 872 ctactccgcttcccaatgcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 931  
DB 841 CTACTCCGCTTCCCTGAGCTGAGCCAGCTGAGGCAAACTCTGCTGCTGCTGCTGCTGCTGCT 900

QY 932 caggaacatctttagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 991  
DB 901 CAGGAACATTTTATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

QY 992 tctgtagcagaaagcaatagcaacaatagcaacaatagcaacaatagcaacaatagcaaca 1051  
DB 961 TCTGTAGCACGAAACATTTAGCAACATATGTCACAAACATATGAGCCCATGAAAA 1020

QY 1052 ccttatttattatagatagcaagaacccctgaataatgcttcttcttcttcttcttcttccca 1111  
DB 1021 CTATTATTTATATATGATGAGGAAACCTGAAATATGCTTCTTCTTCTTCTTCTTCTTCTTCT 1080

QY 1112 atcattaaaaaactgaactactccttagctcgaaggttaagccattcctcagcttagc 1171  
DB 1081 ATCATTTAAAAACGTTAAAACTACTTATAGTCCGAAAGCTTAAAGCATTCTCAGCTTAGC 1140

QY 1172 agtgcagagctgagatctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1231  
DB 1141 AGTGCAGAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

QY 1232 atctctcagagagatgctgagcaaatcttaagagagctgctgctgctgctgctgctgctg 1291  
DB 1201 ATTCTCTAGGAGATGAGGCAAACTATATGGAAGTCTGGAATTAACAGAGCTTAAGANG 1260

QY 1292 calcagacatttccagagagagctagcagacatgcccagctcctcctcctcctcctcctc 1351  
DB 1261 CATCAGACATTTCCAGGAGAGGCTGACATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 1352 ttcactctgttcaatgagagatcttgcagaagaagaatccccaactcagagctcctgagga 1411  
DB 1321 TTTCATCTGTTTATGAGAGATCTTGCAGAGCAAGAAATTCCTCAACCTGAGATGCTGAGGA 1380

QY 1412 ctgtctcttgggaaacactcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1471  
DB 1381 CTGTGCTTTGGGAAACACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

QY 1472 agctgagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1531  
DB 1441 AGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

QY 1532 acaaatcaagctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1591  
DB 1501 ACACATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

QY 1592 accagagcagagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1651  
DB 1561 ACCAGAGCACAGGCTCCTCAGACTTCTGAGCTTCTGAGCTGATGATTAACCTCTAGGAGC 1620

QY 1652 cctggttccagaacctaaagagagcaatcttggttgaatgataaacccttagagac 1711  
DB 1621 CCTGTTTCCAGAACCTTAAGAGAGGACATTTTGGTTGAGATTAACCTCTAGGAGC 1680

QY 1712 catggttagctgcatgctcaggaagaacacccctcctcctcctcctcctcctcctcctcctc 1771  
DB 1681 CATGGGTACTCTCATGCTCAGGAAACATCTCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 1772 gaggtaacttaagatcttctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1831  
DB 1741 GAGGTAATTAAGGATTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

QY 1832 tctatgaatcttctgacagagagagagctggtgatactcctgacagtaataatgata 1891  
DB 1801 TTATTTGAATTAATTTTGCAGGAGAGAGACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1860

QY	1892	catgtgaacagagtttaacaaacacctggttatttttgccttlttattaaaccttg	1951
Db	1861	CAATGTGAACGAGATTACCAACACACCTCTGTGTGTTTGGCTTTTGTATTATACACTTG	1920
QY	1952	ggacaataattttaaattatatacatgcaagagacatcgacgcagagaagactaaagagacttg	2011
Db	1921	GGACAAATTTTTTMAATTTATATCAATGCAGAACTGCAAGCCGAGAGAAAGCTPMAAAGACTTG	1980
QY	2012	ccctctgccacacagcagctggtagagcctgaactcaaacaccagttctcatctcacctca	2071
Db	1981	CCCTCGCCCAACAGCAGCATGTGTAGAGCCTGACACTCAAAACCAGTCTCATCTCACCTCA	2040
QY	2072	gggggctgcttctcccatctgctatgtgccttaagtgatggtgagactaaagaaatgaagt	2131
Db	2041	GGGGGCTGCTTCCCACTGCCTGTATTGTCTTAAAGTGATGGGTACTAGCAATGAAGCT	2100
QY	2132	aattctctaggaagaatgataccaatttcccttctccacatccctcttlttctccaacc	2191
Db	2101	AATTCTCTAGGAAGCATGCCAATATTCCTATCTCCACCTCTCTTTTCTCTCCACCC	2160
QY	2192	ctcccccatagcccccatatataatgccaacatctcccaagaagcttgc	2240
Db	2161	CTCCCCCATAGCCCCCATATATATAGCCCAATCTCCCAAAAGCTTTGC	2209

RESULT	2
LOCUS	AC098935
DEFINITION	AC098935 Homo sapiens chromosome 1 clone RP11-564A8, WORKING DRAFT SEQUENCE.
ACCESSION	AC098935
VERSION	2 unsorted pieces.
KEYWORDS	AC098935 1.GI:16751905
SOURCE	HTG; HTS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFTN. human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 193317) Direct Submission Unpublished 2 (bases 1 to 193317) Direct Submission Submitted (06-NOV-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA On Nov 6, 2001 this sequence version replaced gi:10277966.
JOURNAL	
COMMENT	

```

Center: University of Washington genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: SC
-----
Project Information
Center Project name: chr-1
Center clone name: RP11-564A8 (sc0742)
-----
Summary statistics
Sequencing vector: plasmid; 39% of reads
Sequencing vector: plasmid; 108757; 61% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 132768 bases at least Q40
Consensus quality: 131102 bases at least Q30
Consensus quality: 193186 bases at least Q20
Insert size: 193217; sum-of-ctrls
Insert coverage: 10..0x in Q20 bases; sum-of-ctrls

```

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 56155: contig of 56155 bp in length  
\* 56156 56255: gap of unknown length  
\* 56256 19317: contig of 137062 bp in length.  
\* Location/Qualifiers  
\* 18292..

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-564A8"
/clone_1b="RP11 human BAC library 11"
1..56155
/misc_feature
/note="assembly_name:Contig81"
56256..193317
/misc_feature
/note="assembly_name:Contig82"
56256..193317
BASE COUNT 53599 a 42586 c 43895 g 53137 t 100 others
ORIGIN

```

Query Match	96.5%;	Score 2162.6;	DB 2;	Length 193317;
Best Local Similarity	98.7%;	Pred. No. 0;		
Matches 2207; Conservative	0;	Mismatches 9;	Indels 19;	Gaps 2

OY	25	gacgcagatcacacctttgaaacccagagctctgcctgctcccaaacctctgacccaataacag	84
Db	103494	GGCTCGATCACCTTTTGAAACCCAGAGCTGCGCTGCTCCAAAGCTGTGACTCATTAAC	103553
OY	85	ttctcaactgtagtctgggccaaggtccctcagttctctctctcgaacctctcctc	144
Db	103554	TTCTCAACTGTAGTGTGGGCCAAGGGTCTCTGAGTCTCTCTTCACTTCTCTTGAA	103613
OY	145	ataatgcataataagcaccatcgagagctgagagccagagcaatgcttgcctgaactac	204
Db	103614	ATAATGCTATATTAAGCTCATCGAGAGCTGAGGCCCGACATGTTTGGCTGA	103673
OY	205	catgtatataagatctctctctcctcaagacagatgagatactcaagatacccaagtg	264
Db	103674	CATGTTATATATATTCCTTCCTCAACACAGAGACTCTACACATCCCAAGTGTAC	103733
OY	265	aggccagcccaagagtgatataccataagcccaatgagcctctgttcagagctgccc	324
Db	103734	AGGCCAGCCAAAGGTGATCATAGACCTGATGCTGTCTCGAGCTGCCCCTTAA	103793
OY	325	catccacactctgcgcctctcccccgcctatactgcagaaagtagtcaagataccag	384
Db	103794	CATCCACACTCTGCGCCTCCGCCCTATGCGAGACAGTAGTCAAGATTTCAGCTG	103853
OY	385	ctggggggtctacttttccctctcagaactccctggtttagctgtctccctgcctcc	444
Db	103854	CTGGGGGCTCTATTTCCCTCTCAACCTCTCGCTTTACTCTGTCTCGCCCTCCAC	103913
OY	445	latatcccaagacactctcaactctgctctctctctctcgaatcaactgcctctg	504
Db	103914	TATATCTCTCACACACTCTCACTGGTCTCTTCTTGCTCATCATCTGCTTGACAT	103973
OY	505	ttactcctaagttagttagtaagggtctctctggtataatgcctaaatcaactatg	564
Db	103974	TTATCTCTATAGTATAGTTAATTAGGGGGTTCTTGGTATATCCCTTAATCCAT	104033
OY	565	gggagagctgggggagaagaagctgcgcctctgagagctctgaactctcgaaggt	624
Db	104034	GGGAGAGCTGGGGGAAGAGAGTGGCCTGTGGGGCTGTGCTCACTTCTTGAGAG	104093
OY	625	gggcctctcaggaacaaagatctcaagctctgltgagcagctatagccaagcagct	684
Db	104094	GGGCTCTCCAGGAACAAAGAGATTAGGCTGTGTGCACCTATTAGCCAAGCACT	104153
OY	685	cagagattgcgaagaagatcttctgtcttaagaanaataacaaactgaatagat	744
Db	104154	CAGGATTTGCANAAGAGATTTTGTCTTTTGCTTAAABAAATTAACACACTAG	104213





## COMMENT

On May 4, 2000 this sequence version replaced gi:7143452.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W18R

Web site: <http://www-seq.wi.mit.edu>

Contact: [submissions@genome.wi.mit.edu](mailto:submissions@genome.wi.mit.edu)

Project Information

Center project name: 15244

Center clone name: 462.N.18

## Summary Statistics

Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 194690 bases at least Q40  
Consensus quality: 207622 bases at least Q30  
Consensus quality: 217061 bases at least Q20  
Insert size: 200000; agarose-ftp  
Insert size: 225302; sum-of-coverage  
Quality coverage: 4.5 in Q20 bases; agarose-ftp  
Quality coverage: 4.0 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 41 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      1108: contig of 1108 bp in length
*      1109 1208: gap of 100 bp
*      1209 2504: contig of 1296 bp in length
*      2505 2604: gap of 100 bp
*      2605 3688: contig of 1084 bp in length
*      3689 3788: gap of 100 bp
*      3789 4962: contig of 1174 bp in length
*      4963 5062: gap of 100 bp
*      5063 6375: contig of 1313 bp in length
*      6376 6475: gap of 100 bp
*      6476 7673: contig of 1198 bp in length
*      7674 7773: gap of 100 bp
*      7774 9425: contig of 1652 bp in length
*      9426 9525: gap of 100 bp
*      9526 10997: contig of 1472 bp in length
*      10998 11097: gap of 100 bp
*      11098 12740: contig of 1643 bp in length
*      12741 14467: contig of 1627 bp in length
*      14468 14567: gap of 100 bp
*      14568 16980: contig of 2413 bp in length
*      16981 17080: gap of 100 bp
*      17081 19176: contig of 2096 bp in length
*      19177 19276: gap of 100 bp
*      19277 21882: contig of 2606 bp in length
*      21883 21982: gap of 100 bp
*      21983 23826: contig of 1844 bp in length
*      23827 23926: gap of 100 bp
*      23927 26561: contig of 2635 bp in length
*      26562 26661: gap of 100 bp
*      26662 28846: contig of 2185 bp in length
*      28847 28946: gap of 100 bp
*      28947 32495: contig of 3549 bp in length
*      32496 32595: gap of 100 bp
*      32596 34950: contig of 2355 bp in length
*      34951 35050: gap of 100 bp
*      35051 38193: contig of 3143 bp in length
*      38194 38293: gap of 100 bp
*      38294 41862: contig of 3569 bp in length
*      41863 41962: gap of 100 bp
*      41963 45898: contig of 3936 bp in length

```

## FEATURES

source

```

*      45899 45998: gap of 100 bp
*      45999 50904: contig of 4906 bp in length
*      50905 51004: gap of 100 bp
*      51005 55346: contig of 4342 bp in length
*      55347 55446: gap of 100 bp
*      55447 61336: contig of 5890 bp in length
*      61337 61436: gap of 100 bp
*      61437 65806: contig of 4370 bp in length
*      65807 65906: gap of 100 bp
*      65907 71648: contig of 5742 bp in length
*      71649 71748: gap of 100 bp
*      71749 74886: contig of 3138 bp in length
*      74887 74986: gap of 100 bp
*      74987 79878: contig of 4892 bp in length
*      79879 79978: gap of 100 bp
*      79979 84280: contig of 4302 bp in length
*      84281 84380: gap of 100 bp
*      84381 89873: contig of 5493 bp in length
*      89874 89973: gap of 100 bp
*      89974 97168: contig of 7195 bp in length
*      97169 97268: gap of 100 bp
*      97269 101908: contig of 4638 bp in length
*      101909 102006: gap of 100 bp
*      102007 109439: contig of 7433 bp in length
*      109440 109539: gap of 100 bp
*      109540 117592: contig of 8053 bp in length
*      117593 117692: gap of 100 bp
*      117693 126037: contig of 8345 bp in length
*      126038 126137: gap of 100 bp
*      126138 1235140: contig of 9003 bp in length
*      135141 135240: gap of 100 bp
*      135241 146384: contig of 11144 bp in length
*      146385 146484: gap of 100 bp
*      146485 160352: contig of 13868 bp in length
*      160353 160452: gap of 100 bp
*      160453 173458: contig of 13006 bp in length
*      173459 173558: gap of 100 bp
*      173559 191605: contig of 18477 bp in length
*      191606 191705: gap of 100 bp
*      191706 229302: contig of 37597 bp in length.
Location/Qualifiers
1. .229302
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-462N18"
/clone_id="RP11-462N18 Human Male BAC"
1. .1108
/note="assembly_fragment"
1209. .2504
/note="assembly_fragment"
2605. .3688
/note="assembly_fragment"
3789. .4962
/note="assembly_fragment"
5063. .6375
/note="assembly_fragment"
6476. .7673
/note="assembly_fragment"
7774. .9425
/note="assembly_fragment"
9526. .10997
/note="assembly_fragment"
11098. .12740
/note="assembly_fragment"
12841. .14467
/note="assembly_fragment"
14568. .16980
/note="assembly_fragment"
17081. .19176
/note="assembly_fragment"
19277. .21882
/note="assembly_fragment"
21983. .23826
/note="assembly_fragment"
23927. .26561
/note="assembly_fragment"
26662. .28846
/note="assembly_fragment"
28847. .32495
/note="assembly_fragment"
32496. .34950
/note="assembly_fragment"
34951. .38193
/note="assembly_fragment"
38194. .38293
/note="assembly_fragment"
38294. .41862
/note="assembly_fragment"
41863. .45898
/note="assembly_fragment"

```

Query Match Best Local Similarity 98.7%; Pred. No. 0; Mismatches 2207; Conservative 0; Indels 19; Gaps 2;

misc\_feature /note="assembly\_fragment"  
23827..26561  
/note="assembly\_fragment"  
misc\_feature 26562..28846  
/note="assembly\_fragment"  
misc\_feature 28947..32495  
/note="assembly\_fragment"  
misc\_feature 32596..34950  
/note="assembly\_fragment"  
misc\_feature 35051..38193  
/note="assembly\_fragment"  
misc\_feature 38294..41862  
/note="assembly\_fragment"  
misc\_feature 41963..45898  
/note="assembly\_fragment"

Query 25 gactgaacacatttgaaacacagctgctgcctcccaagctgtactataactaga 84  
Db 158375 GACTGCATGATCACTTTGAAACCCAGGCTGCTGCTGCCCAAACTGTACTCATTAACAGA 158316  
Query 85 ttctcaacagatgttgaggaaggttctcagggttctctcctgacctctctcgaagta 144  
Db 158315 TTCTCAACTGATGTTGGGCCAAGGTTCTTCAAGTTCTCTCTTGACCTTCTTGAAATA 158256  
Query 145 ataatgtaatgaatgaatcagagctgagagcccaagacacatgttgcctgaactac 204  
Db 158255 ATAAATGCTATGATTAAGCTCATCCGAGCGTACGAGCCAGCCACATCTTTCCTGAACATATC 158196  
Query 205 caagttaatgatctctctctcagacagagtgagtaactcaacagatcccaagctgaacct 264  
Db 158195 CATGTATATGATTCCTCTCTCTCAGACAGAGTACGACACGACATCCCAAGGTGTAACCTG 158136  
Query 265 aggcacagcaagctgatacagacagctcagctcctgttccagcctgccttaacagct 324  
Db 158135 AGGCACACCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 158076  
Query 325 catccacacagctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 384  
Db 158075 CATCCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 158016  
Query 385 ctggggagctcaatttccctcctcagctcctcctcctcctcctcctcctcctcctcctc 444  
Db 158015 CTGGGGAGCTCAATTTCCCTCCTCAGCTTCTGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTG 157956  
Query 445 taatacaccagac 504  
Db 157955 TAAATACACGACACTCTCAACCTGCTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157896  
Query 505 ttatctcaatagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 564  
Db 157895 TTATCTCAATAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 157836  
Query 565 gggggaagctggggaagagctgctgctgctgctgctgctgctgctgctgctgctgctg 624  
Db 157835 GGGGGAAGCTGGGGAAG 157776  
Query 625 ggggcccacagaaacaaagatcagagctgctgctgctgctgctgctgctgctgctg 684  
Db 157775 GGGGCCCTCAGGAACAAAGATTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157716  
Query 685 caaggaatgcaaaagagatatttctgctgctgctgctgctgctgctgctgctgctgctg 744  
Db 157715 CAAGGAATGCAAAAGAGATTTTGTGCTTAAGAAATTAACAACTGATGATGATGATGAT 157656  
Query 745 ggaaggaagagagctgctgctgctgctgctgctgctgctgctgctgctgctgct 804  
Db 157655 GGAAGGAG 157596  
Query 805 tcaactcctccttctctcctcctcctcctcctcctcctcctcctcctcctcctcct 864

Db 157596 TCACCTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 157536  
Query 865 caatctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 924  
Db 157535 CAACTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCTCTCCT 157476  
Query 925 aacagagagagagatatttgaagctgctgctgctgctgctgctgctgctgctgctg 984  
Db 157475 AACAGAGAGAGAGATTTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157416  
Query 985 ctctgctcctgtaagcaagaaagcaatlagcaaatatgtaacaaatagtgacccc 1044  
Db 157415 CTCTGCTCTGTAGCAGGAAAGCAATAGCAATATATGTCAACAACATATGTACCCC 157356  
Query 1045 atgaaaacttatattatagatagagaaacatgaaatagatgcttcttcttcttcttct 1104  
Db 157355 ATGAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 157296  
Query 1105 ttccccaatcaataaaaaacgtaaaaaactcctttagtgcgaagtgtaagccattc 1164  
Db 157295 TTCCCAATCATTTAAAAACGTAAAAACTCTTGTAGTTCGCAAGGTTAACGATTCATCA 157236  
Query 1165 gcttagcagtagcagctgctgctgctgctgctgctgctgctgctgctgctgctg 1224  
Db 157235 GCTTAGCAGTAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157176  
Query 1225 aaaaatgattctcgaagagatgtaggaatlaactltaggaagtgctgattaaacag 1284  
Db 157175 AAAATGATTCCTCAGAGAGATGTCGCAAAATCTTATGGAATAATGCTGATTAACAG 157116  
Query 1285 taagaagac 1344  
Db 157115 TAAAGACATACACATTTTCCAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157056  
Query 1345 caatt-ggattcaatctgttcaatgagagatctgcaagaagaatctcctcaaacctag 1403  
Db 157055 CATTAAGATTCATGTTTATGAGAGATCTTTCAGACAAATAATCTCAAACTTAAAG 156996  
Query 1404 tctgagagctgctccttgggaaacacacacacacacacacacacacacacacacacacac 1463  
Db 156995 TCTGAGAGCTGCTCTTGGGAAACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156936  
Query 1464 gaattcagagctgagctgctgctgctgctgctgctgctgctgctgctgctgctgct 1523  
Db 156935 GAATCTAGAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156876  
Query 1524 cagctctgac 1583  
Db 156875 CAGCTTGTACACATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156816  
Query 1584 actgagccacagagac 1643  
Db 156815 ACTGAGCCACAGACACACAGGCTCTCCAGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTG 156756  
Query 1644 ttattctcctcctgctcctcagaaacacacacacacacacacacacacacacacacacac 1703  
Db 156755 TTATCTCTCCTGCTTCCAAACACTTGAAGAGAGGCACTTTTGGTGTAGTATATAAC 156696  
Query 1704 ctaaggaacatggtagctgcatgtaagaagaacacacacacacacacacacacacacac 1763  
Db 156695 CTAGGAGCAATGGGTACTCATGTCAGGAACACCTTCAACCTTCTGAGGCTGAGATGGA 156636  
Query 1764 ttaagaagagagtagctacaggttattcttctgctgctgctgctgctgctgctgctgctg 1823  
Db 156635 TTAAGAAGAGAGTAGTACAGGTATTATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156576  
Query 1824 gggacacattatggaattattctgcaagaagaagagctggtgctgctgctgctgctgct 1883  
Db 156575 GGGCATATATTATTAATTATTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156516  
Query 1884 tgatactacatgctacagagcttcaacagacacacacacacacacacacacacacacacac 1943

Db	156151	TGATTTACATGTGTACAGAGTTTACACAGACCTCTGTGTGTTTWTGCTCTGTTAT	156456
Qy	1944	taacat-----gggaacaatttttaaatcttaacttaagtgaaggaact	1985
Db	156455	TACACTTGCCCTTTGGGGGGAGGTGGACAAATTTTTTTTAAATAATTTATACATGAGAGACT	156396
Qy	1986	gcagcgcaagagaagctaaagaagacttgcacctgcaccacacagcagatgtagagcttgaac	2045
Db	156395	GCAGCCACAGAGAGCTAAAGACTTGCCCTGCCACACAGCCAGTGAGAGCCCTGAAC	156336
Qy	2046	tcaaacccagagctatctcacactgaaggagctgtcttcccaacgcgctgtagtctctaa	2105
Db	156335	TCAACCCAGAGTCTACATCTCACCTTAGGGCCTGCTTTCCCAATGCTGTATTTCTTTAA	156276
Qy	2106	agtgatg99tbgactaggaacaatgaatattctctaggaagaacatgaccaatttccttc	2165
Db	156275	AGTGATGGGGATAGGCATAGTAATTTCTTAGAAGAGCATGACAAATTTCCCTTTC	156216
Qy	2166	tcaacctccctcttcttctccacccccccccaacgaagcccaatataatgcccgaatc	2225
Db	156215	TCCACCTCCCTCTTTTCTCTCCACCCCTCCCCCATCAGCCCCCATATATATGCCCAATC	156156
Qy	2226	tcacacaagcctctgc 2240	
Db	156151	TCCACAAAGCCCTTGC 156141	
RESULT	4		
AC068122		180048 bp	DNA linear HTG 07-JUL-2000
LOCUS	AC068122	Homo sapiens chromosome 1 clone RP11-237C22, WORKING DRAFT	
DEFINITION	SEQUENCE, 28 unordered pieces.		
ACCESSION	AC068122		
VERSION	AC068122.3	GI:79233971	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	Waterston,R.H.		
JOURNAL	Waterston,R.H.		
REFERENCE	The sequence of Homo sapiens clone		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 180048)		
REFERENCE	Waterston,R.H.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (28-APR-2000) Genome Sequencing Center, Washington		
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	MO 63108, USA		
COMMENT	On May 18, 2000 this sequence version replaced gi:7717167.		

```

-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H.NH0237C22
-----
Summary Statistics
Sequencing vector: M13, 100%
Sequencing vector: plasmid, 0%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 158944 bases at least Q40
Consensus quality: 165826 bases at least Q30
Consensus quality: 169627 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 177348; sum-of-configs
Quality coverage: 3.90 in Q20 bases; agarose-fp
Quality coverage: 3.76 in Q20 bases; sum-of-configs
-----

```

	* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence as soon as it is available and the accession number W117109
	* be preserved.
	1 1400: contig of 1400 bp in length
*	1401 1500: gap of unknown length
*	1501 3888: contig of 1888 bp in length
*	33489 3488: gap of unknown length
*	34889 4879: contig of 1391 bp in length
*	4880 6665: gap of unknown length
*	4980 6665: contig of 1986 bp in length
*	6966 7065: gap of unknown length
*	7066 9366: contig of 2301 bp in length
*	9367 11957: contig of 2491 bp in length
*	11958 12057: gap of unknown length
*	14208 14308: contig of 2151 bp in length
*	14309 17022: contig of 2714 bp in length
*	17023 17123: gap of unknown length
*	17123 20061: contig of 2839 bp in length
*	20062 20161: gap of unknown length
*	20162 24025: contig of 3864 bp in length
*	24026 24125: gap of unknown length
*	24126 26718: contig of 2593 bp in length
*	26719 28918: gap of unknown length
*	28919 29036: contig of 2118 bp in length
*	29037 32791: gap of unknown length
*	32792 32791: contig of 3755 bp in length
*	32792 32881: gap of unknown length
*	32882 38895: contig of 6604 bp in length
*	38896 38895: gap of unknown length
*	38996 44306: contig of 5311 bp in length
*	44307 44406: gap of unknown length
*	44407 51589: contig of 7183 bp in length
*	51590 51689: gap of unknown length
*	51690 58591: contig of 6902 bp in length
*	58592 58691: gap of unknown length
*	58692 63442: contig of 4751 bp in length
*	63443 63542: gap of unknown length
*	63543 70554: contig of 7012 bp in length
*	70555 70654: gap of unknown length
*	70655 77794: contig of 7140 bp in length
*	77795 77894: gap of unknown length
*	77895 85841: contig of 7947 bp in length
*	85842 85941: gap of unknown length
*	85942 93327: contig of 7286 bp in length
*	93328 93327: gap of unknown length
*	93328 100871: contig of 7544 bp in length
*	100872 100971: gap of unknown length
*	100972 112141: contig of 11170 bp in length
*	112142 112241: gap of unknown length
*	112242 126358: contig of 14117 bp in length
*	126359 126458: gap of unknown length
*	126459 141767: contig of 15309 bp in length
*	141768 141867: gap of unknown length
*	141868 155927: contig of 14060 bp in length
*	155928 156027: gap of unknown length
*	156028 180048: contig of 24021 bp in length.
FEATURES	
SOURCE	Location/Qualifiers 1. .180048 "Homo sapiens" /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /clone="FPL1-237C22"
BASE COUNT	49152 a 39956 c 40116 g 48114 t 2710 others
ORIGIN	
Query Match	94.0%; Score 2105.4; DB 2; Length 180048
Best Local Similarity	98.3%; Pred. No. 0;



```
OY      2164  tctccacctccctcttcttccacaccccccacatagccacatataatgcccacaa 2223
          |||
Db      82863  TCTCCACCTCCCTCTTTCTCCGACACCCCTCCCGCATGACGCCCATATATATGCCCAA 82922
OY      2224  tctccacaagccttgc 2240
          |||
Db      82923  TCTCCACCAAGCCTTGC 82939

RESULT   5
LOCUS     AY062931                11114 bp    DNA        linear   PRI 26-NOV-2001
DEFINITION Homo sapiens Interleukin 24 (IL24) gene, complete cds.
ACCESSION AY062931
VERSION   AY062931.1  GI:17104816
KEYWORDS
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 11114)
AUTHORS   Rieder,M.J., Armet,T.Z., Carrington,D.P., Chung,M.-W., Lee,K.L.,
           Poel,C.L., Toth,E.J., Yi,Q. and Nickerson,D.A.
TITLE      Direct Submission
JOURNAL    Submitted (13-NOV-2001) Molecular Biotechnology, University of
           Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT    No cite this work please use: SeattleSNPs, NHLBI Program for
           Genomic Applications, UW-FHCRC, Seattle, WA (URL:
           http://pga.mbl.washington.edu).
FEATURES
     source             1..11114
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /rpt_family="MERL_type"
     repeat_region      543..856
                        /rpt_type="dispersed"
     variation          62/
                        /frequency="0.01"
                        /replace="a"
     variation          634
                        /frequency="0.10"
                        /replace="c"
     variation          806
                        /frequency="0.05"
                        /replace="g"
     repeat_region      884..925
                        /rpt_family="MERL_type"
     variation          1619..1620
                        /rpt_type="dispersed"
                        /frequency="0.07"
                        /replace="a"
     repeat_region      1625..1759
                        /rpt_family="SINE"
     variation          1627
                        /rpt_type="dispersed"
                        /frequency="0.02"
                        /replace="c"
     variation          1889
                        /frequency="0.06"
                        /replace="g"
     mRNA              join(1897..2068,2184..2329,3775..3970,4751..4813,
                        5950..6108,6454..6528,7432..8320)
     gene              /gene="IL24"
                        /product="Interleukin 24"
                        1897..8320
     variation          /gene="IL24"
                        2049
                        /gene="IL24"
                        /frequency="0.06"
                        /replace="t"
     CDS              join(2286..2329,3775..3970,4751..4813,5950..6108,
                        6454..6528,7432..7515)
                        /gene="IL24"

/codon_start=1
/product="Interleukin 24"
/protein_id="AAL34146.1"
/db_xref="GI:17104817"
/translation="MAFQQRLOSLMTLARPCPLLATFASQNMVYVLPCLGFTLLMS
QVSGAGODEHFHGPCQVAGVVPQKLMEAFMAVKDPMQADQNTISARLQDEVLQNSD
AESCYLVHRLLEFYLKTVKKNHRTVEVRLKSFSTLANNFVLIVSQDPSQENEMF
SIRDSAHRRPFLFRFAFRQDVEALFRALGEVDLILTMQKRYKI"
2324..2640
/rpt_family="SINE"
/rpt_type="dispersed"
2461
/gene="IL24"
/frequency="0.90"
/replace="a"
2546
/gene="IL24"
/frequency="0.01"
/replace="a"
2625..3110
/gene="IL24"
/note="Region not scanned for variation"
3221
/gene="IL24"
/frequency="0.01"
/replace="c"
3260
/gene="IL24"
/frequency="0.37"
/replace="t"
3292
/gene="IL24"
/frequency="0.01"
/replace="a"
3352
/gene="IL24"
/frequency="0.01"
/replace="t"
3456
/gene="IL24"
/frequency="0.01"
/replace="a"
3517
/gene="IL24"
/frequency="0.02"
/replace="a"
3530
/gene="IL24"
/frequency="0.05"
/replace="a"
3551
/gene="IL24"
/frequency="0.99"
/replace="a"
3623
/gene="IL24"
/frequency="0.37"
/replace="c"
4061
/gene="IL24"
/frequency="0.01"
/replace="a"
4069
/gene="IL24"
/frequency="0.36"
/replace="t"
4082
/gene="IL24"
/frequency="0.02"
/replace="c"
4095..4467
/rpt_family="SINE"
/rpt_type="dispersed"
```



QY 1502 tcagagctcttgagctgcccgtgagctcttgaacatcagctgctgctgcccctgagtg 1561  
DB 1141 tccagagctcttgagctgcccgtgagctcttgaacatcagctgctgctgcccctgagtg 1200  
QY 1562 gttcagagccacacagagccacagactagccacacagagccacagagccctccagctctcgg 1621  
DB 1201 gttcagagccacacagagccacagactagccacacagagccacagagccctccagctctcgg 1260  
QY 1622 gcttgcacagcgtacatcttccatcttcttccgtggttccagaacactaagagagagcaca 1681  
DB 1261 gcttgcacagcgtacatcttccatcttcttccgtggttccagaacactaagagagagcaca 1320  
QY 1682 ttgtgctgagtgagtgatgaacccacagagagagagagagagagagagagagagagagagag 1741  
DB 1321 ttgtgctgagtgagtgatgaacccacagagagagagagagagagagagagagagagagagag 1380  
QY 1742 tcaacttccctgagcctgagtgatgaacagagagagagagagagagagagagagagagagag 1801  
DB 1381 tcaacttccctgagcctgagtgatgaacagagagagagagagagagagagagagagagagag 1440  
QY 1802 gactacagctccacacatgaagaag 1861  
DB 1441 gactacagctccacacatgaagaag 1500  
QY 1862 ggtgacagctcagacagatgaatgattacatgctgacagagagagagagagagagagagagag 1921  
DB 1501 ggtgacagctcagacagatgaatgattacatgctgacagagagagagagagagagagagagag 1560  
QY 1922 tgtgtgttctgcttctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1983  
DB 1561 tgtgtgttctgcttctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1620  
QY 1964 aaatattacatgacag 2023  
DB 1621 aaatattacatgacag 1680  
QY 2024 cagcagcgtgtag 2083  
DB 1681 cagcagcgtgtag 1740  
QY 2084 cccatcgtgtag 2143  
DB 1741 cccatcgtgtag 1800  
QY 2144 aagcagcagcag 2203  
DB 1801 aagcagcagcag 1860  
QY 2204 ccccatatataatgac 2240  
DB 1861 ccccatatataatgac 1897

RESULT 6  
AF276916 AF276916 7025 bp DNA linear PRI 20-DEC-2000  
LOCUS Homo sapiens melanoma differentiation association protein 7 gene,  
DEFINITION complete cds.  
ACCESSION AF276916  
VERSION AF276916.1 GI:11907920  
KEYWORDS  
ORGANISM Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 7025)  
AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.  
JOURNAL Unpublished  
TITLE The human MDA-7 gene  
REFERENCE 2 (bases 1 to 7025)  
AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.  
JOURNAL Direct Submission  
TITLE Submitted (09-JUN-2000) Department of Surgery, University of

Glasgow, Queen Elizabeth Building, Glasgow Royal Infirmary, Glasgow  
G31 2ER, Scotland  
FEATURES  
source  
1..7025  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
join<1804..1847,3285..3480,4261..4323,5460..5618,  
5964..6038,6942..7025)  
/product="melanoma differentiation association protein 7"  
join(1804..1847,3285..3480,4261..4323,5460..5618,  
5964..6038,6942..7025)  
/codon\_start=1  
/product="melanoma differentiation association protein 7"  
/protein\_id="AA041401.1"  
/db\_xref="gi:11907921"  
/translation="MNPORIOSLWTLARPPEPLATASOMHYVLPCLGFTLLMS  
OVSCHOGDFHFGPCOVKIVPKYKNTNNTVEYRKLKSTSLANNEVLIYSOLOPQENEM  
ASCTIVHTLLEFYLKIVPKYKNTNNTVEYRKLKSTSLANNEVLIYSOLOPQENEM  
SIRDSHREFFLEFRAPKOLDEALTLTALGEVDILLTMQRFKL"  
BASE COUNT 1759 a 1668 c 1759 g 1838 t 1 others  
ORIGIN  
Query Match 58.9%; Score 1319.2; DB 9; Length 7025;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1379; Conservative 0; Mismatches 14; Indels 22; Gaps 3;  
QY 848 taticgtcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 907  
DB 1 TATCCTTCCCCCAGCGACATCTCTACTCCGCTTCCAGTGGCTGGACCGAGTTGGGG 60  
QY 908 aaactctctgtaaaagacag 967  
DB 61 AAACCTTCTCTGTAAGAACACAGAGAACTATTTAGCCCTGCTGTCATATGCTCT 120  
QY 968 cagc 1027  
DB 121 CAGTCACACACTCATCTCTGCTCTGTAGCAGAAAGCATTTAGCACAATATGTCA 180  
QY 1028 caaacatagtcagccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1087  
DB 181 CAACCATATGTGACCCCATGAAACTTATTTATATGATAGAGAAACCTGAAATTAAT 240  
QY 1088 gtcttctcttgatcttcttcccaatcaataaagataaagataaagataaagataaagataaagata 1147  
DB 241 GTCTTCTCTTGTGATTTTCCCAATCATTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 300  
QY 1148 aggttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1207  
DB 301 AGGTTAAGCCATTCTCAGCTTAGCAGTGGCAGCTGATTTGGCTGTGTGACCTTAGAGTTG 360  
QY 1208 gccaatccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1267  
DB 361 GCCAATCCCTGATTCACCAAAATGTATTCCTCAGGAGATGTGGGCAAAATATTATGGGCAAAAT 420  
QY 1268 gctgagatlaaacagagatgaagacatcagacatcttccagagagagagagagagagagagagag 1327  
DB 421 GCTGGATTAAACAGATTAAGAACATTCAGACATTTCCAGACGAGGAGTACACATGCTCCAG 480  
QY 1328 ggcctcctaacctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1386  
DB 481 GGCCTCTCACTGACCTCATATTAGATTCATCTGTTTCATGAGAGATCTTCAAGACAGAGA 540  
QY 1387 attcctcaaacctagagctgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1446  
DB 541 ATTCTCAAAACCTAGAGCTGAGAGCTGCTCTTGGGAAACACGCTCTCTTGATGCTCC 600  
QY 1447 tcaactggcagacagtgag 1506  
DB 601 TCACCTGGCAGCATGATGATATCTAGAGCTGAGTCTCTTCTGACCTGAGATAGAGCTCAGA 660  
QY 1507 gctcttgaactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1566



```
Db 661 GCTCTTGAAGTCCCTGGCAGTCTTGACATCATCAGCTGTCTGTGTCCTCGAAGTGGCTCA 720
Qy 1567 gagcacaagcagcagactagaccacagagacagagcctccagcttctcggcttg 1626
Db 721 GAGCCACAGCAGCCAAAGACTAGCCACAGACACAGAGGCTCCAGCTTCTGGGCTTG 780
Qy 1637 tccatgctacatttctcttcttctcgtgttccagaaactaaggagagacatttg 1686
Db 781 TCCATGGGTACATTTCTTATTTCTCTGTGTTCCAGAACCTTAAGAGAGGACATTTTG 840
Qy 1687 gttgagatataaaccttggagacatggtgtgctgcatgtctaggaacaactctctaac 1746
Db 841 GTTAGAGATTATTAACCTTGAGACCATGGGTGCTGCATGTCTAGAAACACTCTCTCAAC 900
Qy 1747 ttcttgcccttgatgataaagagagagactactacaggttatcttctcgtg---gga 1803
Db 901 TTCTTGCCCTGATGAGATTAAAGAGAGACTTACAGGTATTTTTCGCTATTACGA 960
Qy 1804 ctactgtcccaagatgataaggacatcatatgaaattttgacaggaagagacttg 1863
Db 961 CTACTGTCCAGCATGATAGGAGCATCATTTGATTAATTTTGACAGAGAGACTGG 1020
Qy 1864 tgaatgtcacagataaataatgatttataatgtgataagatgataaagacactctgt 1923
Db 1021 TGATGTGTCACAGATTAATTAATGATTATTAATGTGTACAGAGTTTACAGACACTCTGTG 1080
Qy 1924 ttgtttgctcttgtttatttataacttggg-----acaacttttaa 1965
Db 1081 TTGTTTGGCTCTGTTTATTACACTTGGCCTTGGGGAGAGTGGGACAAATTTTAA 1140
Qy 1966 atttataatgtcagagactgcagagcgagagaagaactaagaagacttgccctgccaaca 2025
Db 1141 AATTTATACATGAGAGACTGACGCGACAGAAAGATTAAGACTTGCCCTGCCACACA 1200
Qy 2026 gccagtgtagagcctgaactcaaacccaggtctcatcactcctcagagagctgcttcc 2085
Db 1201 GCCAGGTGAGAGCTGAGACTCAAAACCCAGGCTTATCTCACCCTGAGGGCTGCTTCCC 1260
Qy 2086 catcgtgtatgttctcttaagtgtggtgtagctaggaagtaattctctagaaa 2145
Db 1261 CATCGGTGAAATGCCCTTAAGATGAGGTGAGCAAGTAATTTCTCTAGGAAA 1320
Qy 2146 gaatgccaatttcccttctcgaactcctcttctcgaacccctcccatcagcc 2205
Db 1321 GCATGACAAATTTCCCTTCTCCACCTCCCTCTTTTCTCTCCACCCCTCCCATCAGCC 1380
Qy 2206 cccatatatagcccaaatctccacaagaagccttgc 2240
Db 1381 CCCATATATATGCCCAAAATCTCCACAAAGCCTTGC 1415

RESULT 7
HSA243947 29034 bp DNA linear PRI 09-NOV-2000
LOCUS HSA243947
DEFINITION Homo sapiens VCX-A gene for variably charged X chromosome protein.
ACCESSION AJ243947
VERSION AJ243947.1 GI:5912555
KEYWORDS VCX-A gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 J. X. M., Yen P. H. and Shapiro, L. J.
TITLE Characterization of a low copy repetitive element S232 involved in
the generation of frequent deletions of the distal short arm of the
human X chromosome
Nucleic Acids Res. 20 (5), 1117-1122 (1992)
JOURNAL 92195814
MEDLINE 2 (bases 1 to 29034)
REFERENCE Fukami, M., Kirsch, S., Schiller, S., Richter, A., Benes, V., Franco, B.,
AUTHORS Muroya, K., Rao, E. S., Merker, S., Niesler, B., Ballabio, A., Ansoorge, W.,
Ogata, T. and Rappold, G. A.
```

```

TITLE A member of a gene family on Xp22.3, VCX-A, is deleted in patients
with X-linked nonspecific mental retardation
JOURNAL Am. J. Hum. Genet. 67 (3), 563-573 (2000)
MEDLINE 20395828
REFERENCE 3 (bases 1 to 29034)
AUTHORS Richter, A.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-1999) Richter A., Biochemical Instrumentation,
EMBL, Meyerhofstrasse 1, Heidelberg, 69117, GERMANY
FEATURES
source
1. 29034
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="Xp22.3"
/map="between 40B and DXS1139"
/clone="LNC110"
/sex="male"
/cell_line="Lymphoblastoid"
/clone_lib="Lawrence Livermore X-chromosome specific
library"
/note="Variably charged, X chromosome mRNA on CRI-S232A"
10574..10732
/gene="VCX-A"
/number=1
10574..11947
/gene="VCX-A"
/number=1
10733..11046
/gene="VCX-A"
/number=1
11047..11294
/gene="VCX-A"
/number=2
join(11195..11294,11487..11947)
CDS
11047..11294
/gene="VCX-A"
/codon_start=1
/product="VCX-A protein"
/protein_id="CAB56197.1"
/db_xref="GI:5912556"
/translation="MSPKPRASGPAPKATENAGKRKSSSPSPDPKKTKTKYAKKKA
VRGRGRKKGAATMAATVAPAEASGPAAGSPDQPELPQHELPPEPVSCTGHD
PLDSQSEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVE
PEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVEPEPLDSQSEVE"
Intron
11295..11486
/gene="VCX-A"
/number=2
11487..11947
/gene="VCX-A"
/number=3
11697..11936
/citation=[1]
/rpl_family="RUL"
repeat_region
12097..12297
/citation=[1]
/rpl_family="RUL"
BASE COUNT 9123 a 5623 c 5586 g 8631 t 71 others
ORIGIN
Query Match 4.1%; Score 91.8; DB 9; Length 29034;
Best local similarity 75.5%; Pred. No. 1.5e-14;
Matches 114; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Qy 891 ctgagcccaaggttgcaactcttctgtlaaagacagacagaaactatttagctc 950
Db 4201 CTAAATCAGAGTGTGCAACCTTTCTGTAAAGACACAGACGAAGTATTTAGGCTT 4260
Qy 951 tctgtgccatatagtgtcctagtcacactcactcactcctcctctgttagcaagaagcat 1010
Db 4261 TGAAGACATATAGCTCTTTTACACATACATCTCTGTTGTAGCAAAAAAAGCAGC 4320
Qy 1011 tagcacacatgttcaacaacacatgtgac 1041
Db 4321 CACAGACATATGTAAATTAATGTGGC 4351
```

RESULT 8  
AC108684/c  
LOCUS  
DEFINITION  
SEQUENCE: 4 unordered pieces.  
AC108684  
AC108684.1 GI:18449745  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 174210)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Anarlungue,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbera,J., Benton,U., Blmage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,  
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R.,  
Devalla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,J.F., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louiseged,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., Meleod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokweto,S., Ogih,M., Okunnu,S.,  
Ogunyeye,N., Oviedo,R., Pace,A., Payton,B., Peery,U., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Severy,G.,  
Scheer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,U., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Tamanai,K., Vasquez,L., Vera,Y., Villalón,D., Vlnson,R., Wall,R.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczek,R., Woodson,C., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.  
Direct Submission  
Submitted (31-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: HD0M  
Center clone name: RP11-359020  
Sequencing vector: Plasmid; M77789

Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329first call to  
findPhrapList  
Consensus quality: 170951 bases at least Q40  
Consensus quality: 171318 bases at least Q20  
Estimated insert size: 170999; sum-of-coverage estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 9.9x in Q20 bases; sum-of-coverage estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drift\_data.html).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N. Gaps between the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 76682: contig of 76682 bp in length  
\* 76683 76782: gap of unknown length  
\* 76783 150676: contig of 73894 bp in length  
\* 150677 150776: gap of unknown length  
\* 150777 164461: contig of 13685 bp in length  
\* 164462 164561: gap of unknown length  
\* 164562 174210: contig of 9649 bp in length.  
Location/Qualifiers  
1. 174210  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="Xp"  
/clone="RP11-359020"  
BASE COUNT 52809 a 34893 c 34663 g 51537 t 308 others  
ORIGIN  
Query Match 4.1%; Score 91.8; DB 2; Length 174210;  
Best Local Similarity 75.5%; Pred. No. 1.6e-14;  
Matches 114; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 891 ctggagccaggttgaggcaactcttcgttaagagcagagagacatttagctc 950  
Db 159418 CTAAATCAGAGTTCCTCAACCTTTCCTGTAAGGACGACGAGGATTTTAGGCTT 159359  
QY 951 tgtgtccatgatgttcagtcacactcactcctctcgtctgtgtagcagaaacat 1010  
Db 159358 TGAAGACAAATATGCTCTCTTTTTCACACTACTATCTCTCTCTTTCAGCAAAAAAAGAC 159299  
QY 1011 tagcaacaatatgtcaacaacatatgtgac 1041  
Db 159298 CACGACAAATATGTAATTAATTAATTAATGAGTGGC 159268  
RESULT 9  
AC002090/c  
LOCUS  
DEFINITION  
AC002090  
AC002090.1 GI:2160130  
HTG.  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 137769)  
Hawkins,T.L., Birren,B.W., Fasman,K.H., Nussbaum,C. and Lander,E.S.  
Genomic sequence from Human 17  
Unpublished  
JOURNAL  
REFERENCE  
AUTHORS  
Hawkins,T.L., Reeve,M.P., Christoffersen,A., Birren,B.W.,  
Fasman,K.H., Lander,E.S., McKernan,K., Munro,C., Richardson,P.,  
Barina,N., Chang,A., Cooke,P., Daly,M.J., Forrest,C., Frapp,W.J.,

TITLE	Journal	Submitted (13-MAY-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region	/clone_11b="Research Genetics/Cal Tech CITB9785X-B (plates 1-194)"
REFERENCE	AUTHORS	Hawkins,T.L., Reeve,M.P., Christoffersen,A., Birren,B.W., Fasnman,K.H., Lander,E.S., McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P., Daly,M.J., Devor,K., Forrester,C., Gage,D., Geraghty,K., Hago,B., Harris,K., Huang,J., Hul,L., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Margulis,N., Morrow,J., Nussbaum,C., O'Connor,T., Olotu,A., Peterson,K., Reeve,M.P., Roberts,D., Rollins,G., Stillwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Wilmer,F., Zemtseva,I. and Zody,M.	repeat_region	9. .113
TITLE	Journal	Submitted (29-MAY-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region	/rpt_family="MUTIC"
REFERENCE	AUTHORS	Hawkins,T.L., Birren,B.W., Fasnman,K.H., Nussbaum,C., Lander,E.S., McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P., Daly,M.J., Devor,K., Dewart,K., Forrester,C., Gage,D., Geraghty,K., Hago,B., Harris,K., Huang,J., Hul,L., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Margulis,N., Morrow,J., Nussbaum,C., O'Connor,T., Olotu,A., Peterson,K., Reeve,M.P., Roberts,D., Rollins,G., Stillwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Wilmer,F., Zemtseva,I. and Zody,M.	repeat_region	/rpt_family="Alusx"
TITLE	Journal	Submitted (02-SEP-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region	418. .660
REFERENCE	AUTHORS	Hawkins,T.L., Birren,B.W., Fasnman,K.H., Nussbaum,C., Lander,E.S., McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P., Daly,M.J., Devor,K., Dewart,K., Forrester,C., Gage,D., Geraghty,K., Hago,B., Harris,K., Huang,J., Hul,L., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Margulis,N., Morrow,J., Nussbaum,C., O'Connor,T., Olotu,A., Peterson,K., Reeve,M.P., Roberts,D., Rollins,G., Stillwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Wilmer,F., Zemtseva,I. and Zody,M.	repeat_region	/rpt_family="MLTIC"
TITLE	Journal	Submitted (09-SEP-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region	951. .1021
REFERENCE	AUTHORS	Hawkins,T.L., Birren,B.W., Fasnman,K.H., Nussbaum,C., Lander,E.S., McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P., Daly,M.J., Devor,K., Dewart,K., Forrester,C., Gage,D., Geraghty,K., Hago,B., Harris,K., Huang,J., Hul,L., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Margulis,N., Morrow,J., Nussbaum,C., O'Connor,T., Olotu,A., Peterson,K., Reeve,M.P., Roberts,D., Rollins,G., Stillwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Wilmer,F., Zemtseva,I. and Zody,M.	repeat_region	/rpt_family="(CAT)n"
TITLE	Journal	Submitted (06-JUN-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region	/rpt_family="(CA)n"
COMMENT		On Jun 4, 1997 this sequence version replaced g1:2085774. The Staden databases, finishing information, and all chromatographic files used in the assembly of this clone are available from our anonymous ftp site.	repeat_region	2009. .2282
FEATURES	SOURCE	All repeats were identified using RepeatMasker: Smlt, A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a> . Location/Qualifiers	repeat_region	/rpt_family="LINE2"
		1. .137769	repeat_region	/rpt_family="MIR"
		/organism="Homo sapiens"	repeat_region	/rpt_family="LINE2"
		/db_xref="taxon:9606"	repeat_region	/rpt_family="MIR"
		/chromosome="17"	repeat_region	complement(2709. .2884)
		/clone="4314"	repeat_region	/rpt_family="LINE2"
			repeat_region	complement(3877. .4032)
			repeat_region	/rpt_family="LINE2"
			repeat_region	4630. .4841
			repeat_region	/rpt_family="MIR"
			repeat_region	complement(5998. .6079)
			repeat_region	/rpt_family="(GA)n"
			repeat_region	complement(7528. .7829)
			repeat_region	/rpt_family="Alusx"
			repeat_region	complement(8692. .8879)
			repeat_region	/rpt_family="MIR"
			repeat_region	9321. .9667
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(9688. .9982)
			repeat_region	/rpt_family="Alus9"
			repeat_region	9985. .10210
			repeat_region	/rpt_family="LINE2"
			repeat_region	11052. .11216
			repeat_region	/rpt_family="MIR"
			repeat_region	complement(12296. .12393)
			repeat_region	/rpt_family="(GGGA)n"
			repeat_region	12572. .12644
			repeat_region	/rpt_family="AT-rich"
			repeat_region	12908. .12995
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(14208. .14363)
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(14577. .14686)
			repeat_region	/rpt_family="LINE2"
			repeat_region	14837. .14904
			repeat_region	/rpt_family="(CAGC)n"
			repeat_region	complement(14906. .14963)
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(17782. .17841)
			repeat_region	/rpt_family="AT-rich"
			repeat_region	complement(19230. .19258)
			repeat_region	/rpt_family="AT-rich"
			repeat_region	complement(19446. .19742)
			repeat_region	/rpt_family="AluY"
			repeat_region	19749. .19785
			repeat_region	/rpt_family="AT-rich"
			repeat_region	complement(20413. .20518)
			repeat_region	/rpt_family="MER5B"
			repeat_region	21555. .21649
			repeat_region	/rpt_family="MER5A"
			repeat_region	complement(21855. .21892)
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(22871. .22908)
			repeat_region	/rpt_family="AT-rich"
			repeat_region	complement(23129. .23168)
			repeat_region	/rpt_family="LINE2"
			repeat_region	complement(23296. .23654)
			repeat_region	/rpt_family="LINE2"
			repeat_region	23894. .24398
			repeat_region	/rpt_family="MLT7CA"
			repeat_region	complement(24536. .24675)



				/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."
	misc_feature			complement(175807..176203)
	misc_feature			/note="match: GSS: Em:A0882642"
	misc_feature			176202..176743
	misc_feature			/note="match: GSS: Em:A0374550"
	misc_feature			complement(177911..178307)
	misc_feature			/note="match: GSS: Em:A0131194"
	misc_feature			complement(177997..178301)
	misc_feature			/note="match: GSS: Dm:A0148743"
	misc_feature			178325..178821
	misc_feature			/note="match: GSS: Em:A0420001"
	misc_feature			complement(179120..179217)
	misc_feature			/note="match: GSS: Em:A0545214"
BASE COUNT	54062 a	43323 c	43111 g	52630 t
ORIGIN				
Query Match	4.1%	Score 90.8:	DB 9;	Length 193126;
Best Local Similarity	55.7%;	Pred. No. 3.1e-14;		
Matches 194:	Conservative	0;	Mismatches 152;	Indels 2; Gaps 1;
QY	901	gtctggcaaaccttcctcgttaagaacagacgaactatatttaggcctgctgcat	960	
Db 125379	GGTGACGAACAATTTTCTGTGTAAAGAACAGATAGCGCATATTATTAGGCTTTGGGCCAT	125438		
QY	961	atcgatcagttacaaactatcatcttcgctccgttagcaagaacaatltaagaacat	1020	
Db 125439	ATAATCTCTGTACACAGTCGCACCCTCCGCATTGCAAGTGMAAGCAGCGTAGACAAA	125498		
QY	1021	agtcaacaacatatgagaccocatgaanaaccttatttatatgatgacgaacctga	1080	
Db 125499	ACGTAAATGACAGGGCAATGCGTGCCCCCACTTAACCTTTATGTGGCACCATGAATTTAA	125558		
QY	1081	aataatgtctcttccttgatatttlcccccaatcaaanaacgtaaanaactactcta	1140	
Db 125559	TTTTCATGTGTACCAACTATTTTTCTTTTATTTTCCAMCATTAAAAATGTAAAAAC	125618		
QY	1141	gttcgaagggttaagcaattcttcagcttagcagtggaaggcttgtgatttgcttgacct	1200	
Db 125619	CTTTCTTAACTTGAGAGGTGAGAACAACAAACG--GCATCGGGCTGGATTTGGCTCATGGGCC	125676		
QY	1201	acagctggccaatcccctgattcccaaatgattcccccagagatgag	1248	
Db 125677	ATTGTGTCCGACCCCTGATTGACATCATCCCACTCACAAATGTGG	125724		
RESULT 11				
LOCUS	AC025298	177957 bp	DNA	linear HTG 03-JUN-2000
DEFINITION	Homo sapiens clone RP11-2N21, WORKING DRAFT SEQUENCE, 5 unordered pieces.			
VERSION	AC025298			
KEYWORDS	AC025298.3 GI:8225254			
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.			
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 177957)			
AUTHORS	Barren,B., Linton,L., Nusbaum,C. and Lander,E.			
TITLE	Homo sapiens, clone RP11-2N21			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 177957)			
AUTHORS	Barren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barua,N., Bastien,V., Bedalov,F., Boguslavsky,I., Bouknight,R.B., Brown,A., Burkett,G., Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galante,J., Galinsky,S., Glade,S., Goyette,M., Graham,L., Granda-Baron,M.,			







TITLE  
 Tsaugen, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zheng, L.-H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 192695)  
 Green, E.D.  
 Direct Submission  
 Submitted (27-OCT-2001) NIH Intramural Sequencing Center, 8717  
 Groveport Circle, Gathersburg, MD 20877, USA  
 On Feb 1, 2002 this sequence version replaced gi:15506407.

Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc\_mousehg1.nih.gov  
 -----  
 Project Information  
 Center project name: csb  
 Center clone name: 470J23

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly has been on at least 8x average coverage in 920 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary statistics -----

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: dye-terminator; Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 191421 bases at least Q40  
Consensus quality: 191690 bases at least Q30  
Consensus quality: 191951 bases at least Q20  
Insert size: 182000; agarose-fp  
Insert size: 193295; sum-of-contigs  
Quality coverage: 10.92x in Q20 bases; agarose-fp  
Quality coverage: 10.33x in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitters.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 31657: contig of 31167 bp in length
* 31168 31267: gap of unknown length
* 31268 72728: contig of 41461 bp in length
* 72729 72828: gap of unknown length
* 72829 100593: contig of 27764 bp in length
* 100593 100893: gap of unknown length
* 100893 159516: contig of 58824 bp in length
* 159517 159617: gap of unknown length
* 159617 159693: contig of 33079 bp in length.

```

```

ORFES
source
location/Qualifiers
1..192695
/organism="Papio cynocephalus anubis"
/db_xref="taxon:9555"
/clone="RP41-470J23"
/clone_id="RP41"
1..31167
misc-feature
/note="assembly_fragment
clone_end:SP6
vector_side:left"
1..65403
/misc-feature
/note="clone overlaps with GenBank Accession Number

```

```

misc_feature      AC091227 clone RP41-327F23 (center project name csa)"
                   31268..72728
                   /note="assembly_fragment"
misc_feature      72829..100592
                   /note="assembly_fragment"
misc_feature      100693..159516
                   /note="assembly_fragment"
misc_feature      147682..192655
                   /note="clone overlaps with GenBank Accession Number
AC098698 clone RP41-177A23 (center project name csc)"
                   159617..192655
                   /note="assembly_fragment
clone_end:t7
vector_side:right"
BASE COUNT      58595 a 40608 c 37486 g 55595 t 411 others
ORIGIN

```

Query Match	4.0%;	Score 90;	DB 2;	Length 192695;
Best Local Similarity	62.0%;	Pred. NO. 5.3e-14;		
Matches 160;	Conservative	0;	Mismatches 95;	Indels 3;
			Gaps	1;
Qy 884	cccatcctctgagcccaagcttgggcaaacctctctctgttaagaacccagacggacattt	943		
Db 138335	CCGAGGCTGTGAGATTAAGGTGGCAAAAGTCTCTTTTAAAGAGGCAACAGTAAATATTTT	138414		
Qy 944	taagctctgtgtgcataatgctctcaagtcacaactactactctctgctctgtagacga	1003		
Db 138415	TTGGCTTCTGTGGCCCGGATACAGTCTATTCACATCACTACTCAACTCTGGCATTTAGACCAA	138474		
Qy 1004	aaggaatttagcaacaaatattgtcacacaacatatgtgaccccatgaaactattattat	1063		
Db 138475	AAGGAGTCAACAGCGGTATGAAAATTAATGATAGTACTGTCTTCCAAATTAATTTATTT	138554		
Qy 1064	tgagtaacggaacacttgaaataatgtcttctcttctgatttcttcccaatcaatlaaaaa	1123		
Db 138535	ATGGACCTGGAATATCTGATGTGCATAATATTC---TCACATTTTTCACACCATTTAAAAA	138591		
Qy 1124	cgtaaaactactacttag 1141			
Db 138592	TGTAAAGCTATTCTTTAG 138609			

RESULT	14
LOCUS	AC011060
DEFINITION	Homo sapiens clone RP11-11G16, WORKING DRAFT SEQUENCE, 25 unordered pieces
ACCESSION	AC011060
VERSION	AC011060.5
KEYWORDS	GI:7229784
SOURCE	HTGS: HTGS_PHASE1: HTGS_DRAFT.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Enkayajota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (passes 1 to 181009)
JOURNAL	Barren, B., Linton, L., Nussbaum, C. and Lander, E.
REFERENCE	Homo sapiens, clone RP11-11G16
AUTHORS	Unpublished
	2 (passes 1 to 181009)
	Barren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barne, N., Beckery, R., Boguslavsky, L., Bookhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Develiano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Fuke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lenczky, J., Lien, C., Locke, K., Macdonald, P., Margolis, N., McEwan, P., Mcnair, A., McKernan, K., McLaughlin, J., Melgrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollare, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Thirell, A., Vassiliev, H., Vo, A., Wheeler, T., Wu, X.,

TITLE  
JOURNAL  
COMMENT

Wyman, D., Ye, W.-J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 17, 2000 this sequence version replaced g1:6984405.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L3156

Center clone name: L3156

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap version 0.960731

Consensus quality: 163681 bases at least Q40

Consensus quality: 172073 bases at least Q30

Consensus quality: 175479 bases at least Q20

Insert size: 155000; agarose-fp

Insert size: 178609; sum-of-contigs

Quality coverage: 5.1 in Q20 bases; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1504: contig of 1504 bp in length  
1505 1604: gap of 100 bp  
1605 2791: contig of 1187 bp in length  
2792 2891: gap of 100 bp  
2892 3909: contig of 1018 bp in length  
3910 4009: gap of 100 bp  
4010 5556: contig of 1547 bp in length  
5557 5656: gap of 100 bp  
5657 7482: contig of 1826 bp in length  
7483 7582: gap of 100 bp  
7583 8926: contig of 1344 bp in length  
8927 9026: gap of 100 bp  
9027 10470: contig of 1444 bp in length  
10471 12199: contig of 1629 bp in length  
12200 12299: gap of 100 bp  
12300 14087: contig of 1788 bp in length  
14088 14187: gap of 100 bp  
14188 15707: contig of 1520 bp in length  
15708 15807: gap of 100 bp  
15808 17945: contig of 2138 bp in length  
17946 18045: gap of 100 bp  
18046 20017: contig of 1972 bp in length  
20018 20117: gap of 100 bp  
20118 22115: contig of 1998 bp in length  
22116 22215: gap of 100 bp  
22216 24556: contig of 2341 bp in length  
24557 24656: gap of 100 bp  
24657 27568: contig of 2912 bp in length  
27569 27668: gap of 100 bp  
27669 32527: contig of 4859 bp in length  
32528 32627: gap of 100 bp  
32628 37892: contig of 5265 bp in length  
37893 37992: gap of 100 bp  
37993 41265: contig of 3273 bp in length  
41266 41365: gap of 100 bp  
41366 52400: contig of 11035 bp in length

# FEATURES

52401 52500: gap of 100 bp  
52501 63506: contig of 11006 bp in length  
63507 63606: gap of 100 bp  
63607 79362: contig of 15756 bp in length  
79363 79462: gap of 100 bp  
79463 97023: contig of 17561 bp in length  
97024 97123: gap of 100 bp  
97124 124426: contig of 27303 bp in length  
124427 124526: gap of 100 bp  
124527 149337: contig of 24811 bp in length  
149338 149437: gap of 100 bp  
149438 181009: contig of 31572 bp in length.

Location/Qualifiers  
1. 181009  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-11616"  
/clone\_1bp="RP11-11 Human Male BAC"  
1. 1504  
/note="assembly-fragment"  
1605. 2791  
/note="assembly-fragment"  
2892. 3909  
/note="assembly-fragment"  
4010. 5556  
/note="assembly-fragment"  
5657. 7482  
/note="assembly-fragment"  
7583. 8926  
/note="assembly-fragment"  
9027. 10470  
/note="assembly-fragment"  
10571. 12199  
/note="assembly-fragment"  
12300. 14087  
/note="assembly-fragment"  
14188. 15707  
/note="assembly-fragment"  
15808. 17945  
/note="assembly-fragment"  
18046. 20017  
/note="assembly-fragment"  
20118. 22115  
/note="assembly-fragment"  
22216. 24556  
/note="assembly-fragment"  
24657. 27568  
/note="assembly-fragment"  
27669. 32527  
/note="assembly-fragment"  
32628. 37892  
/note="assembly-fragment"  
37993. 41265  
/note="assembly-fragment"  
41366. 52400  
/note="assembly-fragment"  
52501. 63506  
/note="assembly-fragment"  
63607. 79362  
/note="assembly-fragment"  
79463. 97023  
/note="assembly-fragment"  
97124. 124426  
/note="assembly-fragment"  
124527. 149337  
/note="assembly-fragment"  
149438. 181009  
/note="assembly-fragment"

BASE COUNT 57610 a 36532 c 33173 g 51292 t 2402 others



```

repeat_region /note="match: GSS: Em:AQ094704"
19777..19993
/note="MER58A repeat: matches 2..224 of consensus"
20817..21173
/note="match: GSS: Em:AQ376698"
complement(20830..20939)
/note="match: GSS: Em:AQ229933"
complement(20830..21133)
/note="match: GSS: Em:AQ235822"
complement(20830..21137)
/note="match: GSS: Em:B65379"
21100..21154
/note="11 copies 5 mer aaac 78% conserved"
21183..21410
/note="MER30 repeat: matches 1..230 of consensus"
21647..21773
/note="MIR repeat: matches 119..246 of consensus"
complement(23139..23500)
/note="match: GSS: Em:AQ210496"
23363..23601
/note="MIR repeat: matches 2..256 of consensus"
complement(24372..24810)
/note="match: GSS: Em:AQ178605"
26871..27202
/note="MSTD repeat: matches 1..341 of consensus"
27203..27695
/note="LIPAI0 repeat: matches 5642..6165 of consensus"
27696..27750
/note="MSTD repeat: matches 341..394 of consensus"
27867..27971
/note="MER5A repeat: matches 84..189 of consensus"
27986..28150
/note="MER5A repeat: matches 3..183 of consensus"
28765..29329
/note="match: GSS: Em:B93338"
28777..29151
/note="match: GSS: Em:AQ125799"
29425..29482
/note="25 copies 2 mer ac 87% conserved"
29740..29805
/note="33 copies 2 mer aa 71% conserved"
29763..29807
/note="9 copies 5 mer aaac 82% conserved"
30205..30412
/note="MIR repeat: matches 2..230 of consensus"
30582..30732
/note="MER5A repeat: matches 8..187 of consensus"
30911..31378
/note="match: GSS: Em:AQ546995"
32009..32080
/note="MIR repeat: matches 132..206 of consensus"
33125..32784
/gene="ba31k16.1"
32125..32784
/gene="ba31k16.1"
/note="ba31k16.1 (snRNA binding domain pseudogene)
match: proteins: SW:000567 Tr:09VM69 Tr:09ZRM0 Tr:09VDD1
Tr:094514 Sw:Q21276 Tr:Q9SG37 Tr:Q29158 Sw:Q12460
Tr:080401 Tr:065334 Tr:065335 Tr:Q9USM4"
/codon_start=1
/pseudo
evidence=not_experimental
34666..34977
/note="Alu5g repeat: matches 11..297 of consensus"
36947..36978
/note="16 copies 2 mer aa 87% conserved"
38608..39160
/note="match: GSS: Em:AQ353745"
complement(40339)
/gene="ELAVL2"
complement(40339..43119,43663..43721,51615..51840,
51533..55306,81257..81360,112241..112481,171869..171992)))
/gene="ELAVL2"

```

	/note="match: CDMS: Em:U29808 Em:X85111 Em:M62643 Em:U17602 Em:U12451 Em:LZ6405 Em:U029148 Em:U17597 Em:D26158 Em:S83320 Em:D31953 Em:AF176675 match: ESTS: Em:AA545382 Em:AA081584 Em:A1843112 Em:AA656788 Em:AA692890 Em:AI145457 Em:AV117363 Em:AM524152 Em:AA416369 Em:A1028574 Em:AM491988 Em:AA549705 Em:D61361 Em:AV362071 Em:A1837658 Em:AV094529
Query Match	4.0% Score 89.2; DB 9; Length 183483;
Best Local Similarity	64.6%; Pred. No. 8.g.e-14;
Matches 133; Conservative	0; Mismatches 73; Indels 0; Gaps 0;
QY	891 ctggagcccaagctgttggccaacactcttcctgtaagaaccagacaagaaactatttttgctc 950
DB	10360 CTAGACACAGAGGTCAG6CAAACTTCTGTGAAGAACCAGATGATAAATTATTTAGGCTT 10419
QY	951 tctgtgcataatgagtcctcatcacactactatcttcgcctctgttagcagaagaacgat 1010
DB	10420 TCTGGGCCCATATAGTGCTCTCGTGCACAACACTCACTTCATCTTGCTTTTTAGGGCAAAAAGCAGT 10479
QY	1011 tagcaacaatatgltcaacaacaacatalgtgaaccoccatgaaaaccttatcatatagtatc 1070
DB	10480 TATTAAAAATACATTAATGAATGAGCAGCGCTGTCCCATAAAGAATTTATGTAGCAC 10559
QY	1071 ggaaacctcgaaaataatgctcttctt 1096
DB	10540 TGAAATTTGAATTCATGTAAATTTTT 10565

Search completed: August 16, 2002, 06:44:24  
Job time: 21728 sec



3  
.  
2

.

.

.....







QY 1892 catgtgtacagagttaccacagaccctgtgtgttttgccttcttaccacttg 1951  
 |||||||  
 Db 1861 CATGTGTACAGAGTTACCACAGACCTCGTGTGTGTTCCTTTGTTATACACTTG 1920  
 QY 1952 ggcacaaattttaaattatcacagacagacgcgcagagagaagactgaagacttg 2011  
 |||||||  
 Db 1921 GGCACAAATTTTAAATTTATACATGCACAGACGTCACGAGAGAGCTTAAGAGACTTG 1980  
 QY 2012 cccctgtccacacagcagctgtgtagagcctgtgaactcaaacacaggtctcatcactca 2071  
 |||||||  
 Db 1981 CCCCTGTCCACACAGCAGCATGTGTAGAGCCTGTAACCTCAACACAGTCTCATCTCACTCA 2040  
 QY 2072 gggggtgtcttcccaatgcgtgtatgttcccttaagaatgagtggtgactgaagaaatg 2131  
 |||||||  
 Db 2041 GGGGCTGTCTTCCCATTCCTGTATGTCTTAAAGATGATGCTGACATGCAATGAAGT 2100  
 QY 2132 aatctctagaagaacatcaccaattcccttcccaaccccttcttcccaacc 2191  
 |||||||  
 Db 2101 AATCTCTAGAGAGCATGACCAATTCCTCTTCACACCTCCCTTTTCCCTCACACC 2160  
 QY 2192 ctcccccataagcccccataatataatgcccacaaatctccaaagcctgtgtctgcaaa 2251  
 |||||||  
 Db 2161 CTCCCCCATAGCCCCCATATATATATCCCAAAATCTCCACAAAGCCTTGCTTGCTGCAAA 2220  
 QY 2252 ccttactctgaatgaatcagcagcgtgtgagcg 2286  
 |||||||  
 Db 2221 CCTTACTCTTGAAATGACTTCACAGGCTGGGAGC 2255  
 RESULT 2  
 AC098935 193317 bp DNA linear HTG 06-NOV-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 1 clone RP11-564A8, WORKING DRAFT SEQUENCE,  
 2 unordered pieces.  
 AC098935 AL359089  
 VERSION AC098935.1 GI:16751905  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FUZZTOP; HTGS\_ACTIVEFIN.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 193317)  
 Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
 Direct Submission  
 JOURNAL Unpublished  
 2 (bases 1 to 193317)  
 Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
 Direct Submission  
 REFERENCE Submitted (06-NOV-2001) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 ON Nov 6, 2001 this sequence version replaced gi:10277966.  
 COMMENT  
 ----- Genome Center  
 Center: University of Washington Genome Center  
 Center Code: UWGC  
 Web site: http://www.genome.washington.edu  
 Contact: uwgchgs@u.washington.edu  
 Drafting Center: SC  
 ----- Project Information  
 Center project name: chr-1  
 Center clone name: RP11-564A8 (sc0742)  
 ----- Summary Statistics  
 Sequencing vector: plasmid; 98% of reads  
 Sequencing vector: plasmid; 108752; 61% of reads  
 Chemistry: Dye-terminator ET; 89% of reads  
 Chemistry: Dye-terminator Big Dye; 11% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 192768 bases at least Q40  
 Consensus quality: 193102 bases at least Q40  
 Consensus quality: 193186 bases at least Q20  
 Insert size: 193217; sum-of-contigs  
 Quality coverage: 10.0x in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 56155: contig of 56155 bp in length  
 \* 56156 56255: gap of unknown length  
 \* 56256 193317: contig of 137062 bp in length.  
 FEATURES  
 Location/Qualifiers  
 1..193317  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP11-564A8"  
 /clone\_lib="RP11 human BAC library 11"  
 misc\_feature  
 1..56155  
 /note="assembly\_name:cont1982"  
 56256..193317  
 /note="assembly\_name:cont1982"  
 BASE COUNT 53599 a 42586 c 43895 g 53137 t 100 others  
 ORIGIN  
 Query Match 96.6%; Score 2208.5; DB 2; Length 193317;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 2253; Conservative 0; Mismatches 9; Indels 19; Gaps 2;  
 QY 25 gactcgatcaccttctgaaccagagctcgtcctcccaagcctgtactataactga 84  
 |||||||  
 Db 103494 GCGTCGATGACACCTTTTGAACCCAGAGTGTGCTCCCTCCCAACCTGTATCTCACTGA 103553  
 QY 85 ttctcaactgtatgttgagcacaaggttcctaggtctcctcctgtgacctctctgaatga 144  
 |||||||  
 Db 103554 TTCTCACTGATGTTGGGCCAAGTTCTCTAGTTCTCTCTGACCTTCCTCAAGTA 103613  
 QY 145 ataattctatgataagctatcagaggtgaggtccagagcaatgtttgctgaatc 204  
 |||||||  
 Db 103614 AATAAGCTATGATATAGCTCATGCGAGGCTGAGGCCACATGTTGCTGAACTATG 103673  
 QY 205 catgtatcatatctcctcccaagagctgagctcctcctcctcctcctcctcctcctc 264  
 |||||||  
 Db 103674 CATGTTATATATCTCTCTCCACAGAGTGTAGTACAGATCCAGAGTACCTG 103733  
 QY 265 aggcagcagcagagtgatcagatgacctgacctgtctcagcctgaccttaacagct 324  
 |||||||  
 Db 103734 AGGCCAGCCAAAGGTATATCCATGACCTCATGCTGTTCACAGCTGCTTAAACAGCT 103793  
 QY 325 catccacactcgtcgtccctcccgccatctgacagacagtagctagatctcaagctgc 384  
 |||||||  
 Db 103794 CATCCACCTGCTGCTCCCTCCGCTCATCTGCAGACAGTACGTAGATTTACACTGCC 103853  
 QY 385 ctgggggctcatcttccctcctcagctcctgctcctgctcctgctcctcctcctcctc 444  
 |||||||  
 Db 103854 CTGGGGGCTCATCTTCCCTCCACCTTCGCTTATGAGCTGCTCCCTCCACCTCAC 103913  
 QY 445 tatctac 504  
 |||||||  
 Db 103914 TATTACTCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 103973  
 QY 505 ttatctcatagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 564  
 |||||||  
 Db 103974 TTATCTCATATAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 104033  
 QY 565 gggggag 524  
 |||||||  
 Db 104034 GGGGAG 104093  
 QY 625 gggccctccaggaacaaagatcagagctggtgagcagctatagcagagagagagagag 684  
 |||||||  
 Db 104094 GGGCCCTCCAGGAACAAAGATTGAGGCTGTGAGGCTATAGCCAAAGAGACTCTCGC 104153

OY	685	caggagcttgaaaggagtgatcttctgtcttgcttaagaagaaataaacaacacgagtagagat	744
Db	104154	CAGGAGATTCGAAGAGAGTATTTTGTTCCTTAGAGAAATTAACACACGTAGATGAGAT	104213
OY	745	ggaaggagagggtgctgtgtctgcagagagatctggaaagagctgtccaaagggtgtctctac	804
Db	104214	GGAAGGAGGGGGGTGGTGCCAGAGAAATTTGGAAAGAGCTGCCAAGGGTGTGTTCAC	104273
OY	805	tcactctccctcttctctcactccactgagctcgagggagagtagatccgtctccccaagt	864
Db	104274	TCACTCTCCTCTTCTTTCATCTCCACTGACGTGGAGGAGATTACGTCTCCCAAGT	104333
OY	865	caaatctcctcctccgttttcccaatgacctgagcccaagtttgaggaaacctcttcgttaag	924
Db	104334	CACATTCCTAATCTCCGCTTTCCTCCATGCGCTGGAGCCAGGATTGGAAACCTTCTGTAAAG	104393
OY	925	aaccagacagaaactatctttagagctctgtgtgcatactgtctccagtcacaactcact	984
Db	104394	AACCAGACAGAGAACTATTTTACGCTCTGTGTGCATATGTCTCATCTCCACATCTACTAT	104453
OY	985	ctctgcacctgtatgcacgaaagcaatagacaacatactgcacaaacaaactaigtgacccc	1044
Db	104454	CTCTGCCTCTGTATGCACGAAACCAATTATGCACAAATATGTCAACAACTATGTATGCC	104513
OY	1045	atgaaaaacttatctatctatctgatacgggaaacctgaaataatgctcttcttctgattt	1104
Db	104514	ATCAAAAACCTTATTTATTTATGTATACGAAACCTGAAATAATATGTCTCTTTCATTTT	104573
OY	1105	tccccaacatcatctaaaaaacgtaaaanactctcttagctgcgaaaggttagagcccttcca	1164
Db	104574	TTCGCCAATCATTTAAAAACGTAATAACTCTTAGAGTGCAGGTTTAGCCATTTCTCA	104633
OY	1165	gctctgacgtgtgcaggtctgagatttgctctgtgacctaaagctgtgccaatccctgattccc	1224
Db	104634	GCTTACAGAGTGGAGGCTGGATTTGGCTTGTGACCTTACAGTTGGCCAACTCCGTATTTCC	104693
OY	1225	aaaatgatactccctcagggatgtgtggcaaaactctatctggaaggtgtctgtgaaatgaaggt	1284
Db	104694	AAATATGATTTCCCAAGGATGTGGGCAAAATCTTATGGGAAATGCTGGATTAAACAGAGT	104753
OY	1285	taagaagcatcagaacaattccagaaggttagcaatgtccagggtctcttaactgacct	1344
Db	104754	TAGAAGCATATCAGACATTTCCAGAGCGGCTTACACATGTCCAGGGCTCTCTAACGTGACCT	104813
OY	1345	caatc-gaatcactcgtttctcatgtgagagatctctgaagaacaaagatctcccaactagag	1403
Db	104814	CATTAGGATTCATCTGTTCATATGGAGAGATCTTGCCAAAGACAAATTTCCCAAACTTAGAG	104873
OY	1404	tctgaggaactgtgctcttggaaacaactgtctcgtctgatagtccctcaactgtggacaatgca	1463
Db	104874	TCTGAGGACTGTCTTTTGGGAAACACTGCTCTGTGATGTCCTCCACATGGGACACTGTGTA	104933
OY	1464	gaactctagagctgagtgctctgtctagctcggagagataggtctaaagctctgtgctgcctgg	1523
Db	104934	GAATCTTAGAGCTGAGTGGCTTGCTTAGGTGAGATAGGGTTCAGAGCTCTTCACTGCCCTGG	104993
OY	1524	gagctctgagacaatcaagcgtctcgtgtctccctgaggtgtgtltaagagccaacagagccaag	1583
Db	104994	CAGCTTGTGACATCACTACCTGTCTGTGTCTCCCTGATGCTTCAAGAGCCACACAGGCCAAG	105053
OY	1584	actagccacacagagcacacagagctctcccaagctcttctggagctgttccatctgtctatcttc	1643
Db	105054	ACTAGCCCAACAGACACACAGGCCCTCCAGCTTCTGTGGGCTGTGCATGSCGTACATTTCC	105113
OY	1644	tctatctctccctggcttccagaaccttaaggagagagacatttgtcttgagtgattataacc	1703
Db	105114	TCTATCTCTCTCTGTTTCCAGAACTTAAGAGAGGACCAATTTTGTGTGAGTATTTATAC	105173
OY	1704	cttagggacaactgtgttagctgtcgtlctcagaaacaacacctcaactctccgtccctgtatgga	1763
Db	105174	CTTAGGACCACTGSGTATGCTGCTCATGAGAAACACTTCTCACTTCTGTGGCCGTGATGGA	105233
OY	1764	ttaaaagagaagttactaaagattatcttctgtcgtgtgtgactactgtlcccaagcatgata	1823

Db	105234	TTTAAGGAGAGGACTACTACAGTTATATTCTTCGTGTGGACATCTGTCCAGCATGTGATA	105293
QY	1824	gggcacatcatattgaattatcttttcaggaagagaciggatgctgcacagtaaa	1883
Db	105294	GGGCAATATATTTGAATTTATTTTGGACAGAGAGAGACTGTGTATGTGCACACTAATAA	105353
QY	1884	tgatattacatgltgtaacagagtttaaccagcacctcgtctgtcttttcgtccttgtaat	1943
Db	105354	TGTATTACATGTGTGTACAGAGTTTACCAAGACCTCTGTGTGTTTGGTCCATGTTTAT	105413
QY	1944	tacacct-----gggacaaattttaaatataacatgacagagac	1985
Db	105414	TACACTTGCCCTTGGGGGAGGTGGAGCAAAATTTTTTAAAAATTTTATACATGCGAGAACT	105473
QY	1986	gcagcgacagagaagctlaagagactgtccctgcaccacaagccagctgltagaccgaaac	2045
Db	105474	GCACGTGAGAGAAAGCTTAAAGACCTTGGCCCCCTGCCACACAGCCACTGTGTAAAGAGCTTAAAC	105533
QY	2046	tcaaacccagcttcacttcacactcaagggctgcttccaccatgcgctlatatgctctaa	2105
Db	105534	TCAAAACCAAGTCTCATCTCAACTCAGGGGCTGTTTCCCATCGCTGATTTGTCCTTAA	105593
QY	2106	agctatgggagactggcaatgaagatattctctagaaagatagaacaatttcctttc	2165
Db	105594	AGTATGTGGTACTAGGCAATAGATATATCTCTAGAAAGCATCAACAAATTTCCCTTTC	105653
QY	2166	tccacctccctcttttctccacacctcccccacaaagccccaatalatgatgcccaatc	2225
Db	105654	TCCACCTCCCTTTTCTCTCCACCCCTCCCATGACGCCCATATATATATGCCCCAATC	105713
QY	2226	tccacaagcctgtcgtcgtcccaaaccttaactcttgaaagagacttccaaaggttgagc	2285
Db	105714	TCCACAAGGCTTGCTGTGCTGCCAACAACCTTTACTCTCGAAATGACTTCCACAGGCTGGAC	105773
QY	2286	g 2286	
Db	105774	G 105774	

RESULT	3	229302 bp	DNA	linear	HTG 04-MAY-20
LOCUS	AC023534				
DEFINITION	AC023534 Homo sapiens clone RP11-462N18, WORKING DRAFT SEQUENCE, 41				
ACCESSION	AC023534				
VERSION	AC023534.3	GI:7705140			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 229302)				
TITLE	Birren, B., Linton, L., Nussbaum, C. and Lander, E.				
JOURNAL	Homo sapiens chromosome, clone RP11-462N18				
REFERENCES	Unpublished				
AUTHORS	2 (bases 1 to 229302)				
	Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldridge, D., Barina, N.C., Bedalov, F., Boguski, L., Bourkigal, B., Brown, A., Burkett, G., Campiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArillano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M., Feneclat, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Larcocque, K., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuff, A., McKernan, K., McPherson, R., Meldrum, J., Menais, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T.M., Peterson, K., Pierre, N., Plisak, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rottman, D., Roy, A., Santos, R., Schauer, S.,				

# TITLE JOURNAL COMMENT

Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Subramanian, A., Talamas, J., Testaye, S., Theodose, J., Tittell, A.,  
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and  
Zody, M.

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2000 this sequence version replaced g1:7143452.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: RIBR

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information  
Center project name: L5344

Center clone name: 462\_M18

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 194690 bases at least Q40

Consensus quality: 207622 bases at least Q30

Consensus quality: 217061 bases at least Q20

Insert size: 200000; agarose-gel

Insert size: 225302; sum-of-coverage

Quality coverage: 4.5 in Q20 bases; agarose-gel

Quality coverage: 4.0 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently  
consists of 41 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1108: contig of 1108 bp in length  
\* 1109 1208: gap of 100 bp  
\* 1209 2504: contig of 1296 bp in length  
\* 2505 2604: gap of 100 bp  
\* 2605 3688: contig of 1084 bp in length  
\* 3688 3788: gap of 100 bp  
\* 3789 4962: contig of 1174 bp in length  
\* 4963 5062: gap of 100 bp  
\* 5063 6375: contig of 1313 bp in length  
\* 6376 6475: gap of 100 bp  
\* 6476 7673: contig of 1198 bp in length  
\* 7674 7773: gap of 100 bp  
\* 7774 9425: contig of 1651 bp in length  
\* 9426 9525: gap of 100 bp  
\* 9526 10937: contig of 1472 bp in length  
\* 10938 11097: gap of 100 bp  
\* 11098 12740: contig of 1643 bp in length  
\* 12741 12840: gap of 100 bp  
\* 12841 14467: contig of 1627 bp in length  
\* 14468 14567: gap of 100 bp  
\* 14568 16980: contig of 2413 bp in length  
\* 16981 17080: gap of 100 bp  
\* 17081 19176: contig of 2096 bp in length  
\* 19177 19276: gap of 100 bp  
\* 19277 21882: contig of 2606 bp in length  
\* 21883 21982: gap of 100 bp  
\* 21983 23826: contig of 1844 bp in length  
\* 23827 23926: gap of 100 bp  
\* 23927 26561: contig of 2635 bp in length  
\* 26562 26661: gap of 100 bp  
\* 26662 28846: contig of 2185 bp in length  
\* 28847 28946: gap of 100 bp  
\* 28947 32495: contig of 3548 bp in length

## FEATURES

source

misc\_feature 32496..32595: gap of 100 bp  
misc\_feature 32596..34950: contig of 2355 bp in length  
misc\_feature 34951..35050: gap of 100 bp  
misc\_feature 35051..38193: contig of 3143 bp in length  
misc\_feature 38194..38293: gap of 100 bp  
misc\_feature 38294..41862: contig of 3569 bp in length  
misc\_feature 41863..41962: gap of 100 bp  
misc\_feature 41963..45898: contig of 3936 bp in length  
misc\_feature 45899..45998: gap of 100 bp  
misc\_feature 45999..50904: contig of 4906 bp in length  
misc\_feature 50905..51004: gap of 100 bp  
misc\_feature 51005..55346: contig of 4342 bp in length  
misc\_feature 55347..55446: gap of 100 bp  
misc\_feature 55447..61337: contig of 5890 bp in length  
misc\_feature 61338..61436: gap of 100 bp  
misc\_feature 61437..65806: contig of 4370 bp in length  
misc\_feature 65807..65906: gap of 100 bp  
misc\_feature 65907..71648: contig of 5742 bp in length  
misc\_feature 71649..71748: gap of 100 bp  
misc\_feature 71749..74867: contig of 3138 bp in length  
misc\_feature 74868..74966: gap of 100 bp  
misc\_feature 74967..79878: contig of 4892 bp in length  
misc\_feature 79879..79978: gap of 100 bp  
misc\_feature 79979..84280: contig of 4302 bp in length  
misc\_feature 84281..84380: gap of 100 bp  
misc\_feature 84381..89873: contig of 5493 bp in length  
misc\_feature 89874..89973: gap of 100 bp  
misc\_feature 89974..97168: contig of 7195 bp in length  
misc\_feature 97169..97268: gap of 100 bp  
misc\_feature 97269..101906: contig of 4638 bp in length  
misc\_feature 101907..102006: gap of 100 bp  
misc\_feature 102007..109439: contig of 7433 bp in length  
misc\_feature 109440..109539: gap of 100 bp  
misc\_feature 109540..117592: contig of 8053 bp in length  
misc\_feature 117593..117692: gap of 100 bp  
misc\_feature 117693..126037: contig of 8345 bp in length  
misc\_feature 126038..126137: gap of 100 bp  
misc\_feature 126138..125140: contig of 9003 bp in length  
misc\_feature 135141..135240: gap of 100 bp  
misc\_feature 135241..146384: contig of 11144 bp in length  
misc\_feature 146385..146484: gap of 100 bp  
misc\_feature 146485..160352: contig of 13868 bp in length  
misc\_feature 160353..160452: gap of 100 bp  
misc\_feature 160453..173458: contig of 13006 bp in length  
misc\_feature 173459..173558: gap of 100 bp  
misc\_feature 173559..191605: contig of 18047 bp in length  
misc\_feature 191606..229302: gap of 100 bp  
misc\_feature 229303..229302: contig of 37397 bp in length.  
Location/Qualifiers  
1..229302  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="RP11-462N18"  
/clone\_id="RPCT-11 Human Male BAC"  
1..1108  
/note="assembly-fragment"  
1209..2504  
/note="assembly-fragment"  
2505..3688  
/note="assembly-fragment"  
3689..4962  
/note="assembly-fragment"  
4963..6375  
/note="assembly-fragment"  
6376..7673  
/note="assembly-fragment"  
7674..9425  
/note="assembly-fragment"  
9426..10937  
/note="assembly-fragment"  
10938..12740  
/note="assembly-fragment"  
12741..12840  
/note="assembly-fragment"  
12841..14467  
/note="assembly-fragment"

```

/notes="assembly-fragment"
misc_feature 14568..16980
/notes="assembly-fragment"
misc_feature 17081..19176
/notes="assembly-fragment"
misc_feature 19377..21882
/notes="assembly-fragment"
misc_feature 21883..23826
/notes="assembly-fragment"
misc_feature 23927..26561
/notes="assembly-fragment"
misc_feature 26662..28846
/notes="assembly-fragment"
misc_feature 28947..32495
/notes="assembly-fragment"
misc_feature 32596..34950
/notes="assembly-fragment"
misc_feature 35051..38193
/notes="assembly-fragment"
misc_feature 38294..41862
/notes="assembly-fragment"
misc_feature 41963..45898
/notes="assembly-fragment"

Query Match 96.6%; Score 2208.6; DB 2: Length 229302;
Best local similarity 98.8%; Pred No. 0;
Matches 2253; Conservative 0; Mismatches 9; Indels 19; Gaps 2;

Oy 25 gactcgtacaccccttggaacccagctcgtcgtcccaaaagcttgatcataactaga 84
DB 158375 GGCCTCGATCAGCCTTTTGAACCCAGGCTGCTCCGCTCCCAACCTTGATCATACAGAGA 158316
Oy 85 tctcacaactgactgtgagccaaagttccclaggttctctctctgacccctctcgaagta 144
DB 158315 TTCTCACTGATGTGGGGCAAGGTCTAGGTTCTCTCTGACCTTCTCTGAAGTA 158256
Oy 145 ataattgactaagatagctacgagagctgagcccaagacacatgttctgctgaagatc 204
DB 158325 ATATGCTATGATTAAGTCTATCGAGAGCTGAGGCCAGGACACATTTTGGCTGAACTATC 158196
Oy 205 cabgttatagatctctctctcagacagagtgagctacatcagatcccaagtgtaacctg 264
DB 158195 CATGTTATGATTAATCTCTCTCAAGACAGTGAAGTACACAGATCCAGATCCAGGTGTACCTG 158136
Oy 265 aggcacagccaaagtgtatcatagctcaatgctcgtctcagcctgaccttaaacgct 324
DB 158135 AGGCACAGCAAGGTGTATCATGACTCATGCTCTGTTCAAGCTGTGCTTAAACAGCT 158076
Oy 325 catccacacgtcgtcgtcctcccgctatctgcagacagtagtctagaattcagctgctc 384
DB 158075 CATCCACACTGCTGCTGCTCCGCTATCTCAAGACAGTGAAGTATTCAGCTGCC 158016
Oy 385 ctggggagctcaattctcctctcagctctcgtctctagctgtctcctcgtcccaactcagc 444
DB 158015 CTGGGGGCTCATTTTCCCTCAGCTTCCTGTTAGGCTGCTCCTGCTCCCACTCAGC 157956
Oy 445 tatatcccccagcactctcaactgtctctctctctctctctctctctctctctctctctc 504
DB 157955 TATTAATCTGACACTCTCACTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 157896
Oy 505 ttatctcatagtagttagttaggttctctgttaaigtccctaaatccatctggtggaag 564
DB 157895 TTATCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 157836
Oy 565 ggaggagtagtgggagagagagtagctgtgaggtctgtgctactctctggaaggtgaagctc 624
DB 157835 GGGGGAATGGGGGAAGAGATGCTGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157776
Oy 625 gggccctccaggaacaaagatctcagctgtgtgagctatagccaagcagactgctgctc 684
DB 157775 GGGCCCTCCAGGAACAAAGATTCAGGCTGTGGGCTGATAGCAAGCAACAGCTGCTGCTGCTG 157716
Oy 685 cagggtctgcaagagtagtctgtctgcttaagaataaacaacactgtagcttagat 744

|||||
DB 157715 CAGGAGATTCGAAGAGATATTTGTTGCTTAAGAAATTAACAACACTGATGATGAT 157656
Oy 745 ggaaggagagagagtagtctccagagagatctgggaaggtctgccaaggtgtctctac 804
DB 157655 GGAGGAGAGGGGGTGTGTGTCCAGAGAGATTTGGAGAGAGTCTCCCAAGGCTGTCTTAC 157596
Oy 805 tcaactccctctctctctctctctctctctctctctctctctctctctctctctctct 864
DB 157595 TTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 157536
Oy 865 cacatctcactcccgcttcccaatgcttgagccaggtgtgggcaaacctctcctgaag 924
DB 157535 CACATTCCTACTCCCGTTTCCATGCTGAGCCAGTGGGCAAACTCTCTGTAAG 157476
Oy 925 aaccagagcaggaactctttagctctgtgtgccaatagttcagttcaaacactactcat 984
DB 157475 AACCAAGAGAGAACTATTTAGGCTGTGTGCTCATGTGCTCATGCTCAACTACTCAT 157416
Oy 985 ctctgctctctgtagcaagaaagcaatagcaaaatagtaacaacataatgtgacccc 1044
DB 157415 CTCTGCTCTGTAGCAAGAGCAATTAACACATATGCTCAACAAACATATGTGACCC 157356
Oy 1045 atgaaaccttatattatagatagaggaacctaataatgtctctctctctctctctct 1104
DB 157355 ATGAACCTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157296
Oy 1105 ttcaccaatcaataaagaataaactactcttagtctgcaaggttaagcaattctca 1164
DB 157295 TTCCCAATCATTAATAAAGATTAATAAAGTATTAATGCTTCAAGGTAAAGCATCTCA 157236
Oy 1165 gcttagagagtagagagctgactgtgtgtgtaacctgaagcttgagcaatccctgactccc 1224
DB 157235 GCTTAGAGAGTGGAGGCTGATTTGGCTGTGACTTACAGTTGGCCAACTCCGATTTCC 157176
Oy 1225 aaatgfatctcccaagagatgtgggcaaatctcttggaaggtctctgattaaacaggt 1284
DB 157175 AAATGATTTCTCTCAGAGATGTGGCAATCTTATGGAATCTGATGATTAACAGAGT 157116
Oy 1285 taagaagcatcaagacatctccagagcgttagcaatctgcaaggtctctaaactgact 1344
DB 157115 TAAAGACATCAAGACATTTCCAGAGAGGCTGTGACATGCTCCAGGCTCTCAACTGACT 157056
Oy 1345 catc-aggatcatctgtctcaatgagagctgtgcaagaacagatctcctaaacctagag 1403
DB 157055 CATTAAGATTCATCTGTTCATGAGAGATCTTGCAAGACAAAGATTTCTCAAACTAGAG 156996
Oy 1404 tctgagagctgtgtcttgagaaacacgtctgtctgtctgctcctactctggaatgtgta 1463
DB 156995 TCTGAGGACTGTCTTGGGAAACACTGCTGTGATGCTTCACTGAGGACATGTGA 156936
Oy 1464 gaactctagagctgtagctctgtctagctgagataggtctcaagctctctgactgacctg 1523
DB 156935 GAATCTAGACTGAGTGTGCTGTGCTGAGATGAGGTGACAGCTTTCATGCTGCTG 156876
Oy 1524 cagctctgaaacatcaagctgtctgtgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgt 1583
DB 156875 CAGCTTGACACATCAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156816
Oy 1584 actagccacacagagacagacagcctccagctctctgtgctgtgtgtgtgtgtgtgtgtgtgt 1643
DB 156815 ACTAGCCACACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 156756
Oy 1644 ttatctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1703
DB 156755 TTATTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156696
Oy 1704 ctgaagaccaaaggtgtagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1763
DB 156695 CTAGGACCAATGGGTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156636
Oy 1764 ttaagagagagtagttaggttattctctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1823

```

Db 15635 TTAAGGAGAGCTACTACAGCTATTTCTTCGCTGTGACACTACGTCCAGACGAAATA 156376  
 QY 1824 GGGGcattatgattatatttgacaggaagagactgtaglgtcgcacagiatata 1883  
 Db 156575 GGGCATTATTATTAATTTTGCACAGAGAGACTGCTGATCTGCACAGTAATA 156516  
 QY 1884 tgaattatattgtacagagatttaccagagactcgtgtgttttttccttgattat 1943  
 Db 156515 TGATTTACATGTGTACAGAGTTTACAGACCTCTGTGTGTGTGTGTGTGTGTGTAT 156456  
 QY 1944 tacactt-----gggacaaattttaaattatatacagagagact 1985  
 Db 156455 TACACTTCCCTTGGGGAGAGTGGACAAATTTTAAATTTATACATCAGAGACT 156396  
 QY 1986 gcaagcagaagaagctaaagagactgccccgcgcacacagcagagtgtagagaccgtaac 2045  
 Db 156395 GCGAGCCAGAGAGCTAGAGAGACTTCCGCCACACAGCCAGTGTAGAGCCCTGAAC 156336  
 QY 2046 tcaaaccaagtcacatcaccacagagagctgtcttcccaatcgctgtatgtccttaa 2105  
 Db 156335 TCAAAACCCAGGTCATCTCACCTCAGGGGCTGCTTCCCATTCGCTGATGTCTCTTA 156276  
 QY 2106 agtgaatgagtgactgagcaatgaatlaattctcagaagaagcagacattcccttc 2165  
 Db 156275 AGTGATGGTGACTAGGCAAGAGTATTTCTTAGGAAAGCATGACCAATTTCCCTTTC 156216  
 QY 2166 tcaactccctcttttctcaccctccctcccaatcagcccatatataatgcccaatc 2225  
 Db 156215 TCCACCTCCCTTTTCTTCTCCACCCCTCCCATGAGCCCATATATATGCCCCAATC 156156  
 QY 2226 tccacaagagctgtctgctcgaacacttactcgaatctactccagagctgtgac 2285  
 Db 156155 TCCACAAGAGCTGTGCTGCTGCAAACTTACTTCTGAATGACTTCCAGGCTGGAC 156096  
 QY 2286 g 2286  
 Db 156095 G 156095  
 RESULT 4  
 AC068122 180048 bp DNA linear HTG 07-JUL-2000  
 LOCUS Homo sapiens chromosome 1 clone RP11-237C22, WORKING DRAFT  
 DEFINITION AC068122  
 ACCESSION AC068122.3 GI:7923971  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 180048)  
 TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 REFERENCE Submitted (28-APR-2000) Genome Sequencing Center, Washington  
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 COMMENT On May 18, 2000 this sequence version replaced gi:7717167.  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 ----- Project Information -----  
 Center project name: H\_NH0237C22  
 ----- Summary Statistics -----  
 Sequencing vector: M13, 100%  
 Sequencing vector: Plasmid, 0%  
 Chemistry: Dye-primer ET, 100% of reads

Chemistry: Dye-terminator Big Dye, 0% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 158944 bases at least Q40  
 Consensus quality: 155826 bases at least Q30  
 Consensus quality: 169627 bases at least Q20  
 Insert size: 167000; agarose-fp  
 Quality coverage: 3.90 in Q20 bases; agarose-fp  
 Quality coverage: 3.76 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 28 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 1 1400: contig of 1400 bp in length  
 1401 1500: gap of unknown length  
 1501 3388: contig of 1888 bp in length  
 3389 3488: gap of unknown length  
 3489 4879: contig of 1391 bp in length  
 4880 4979: gap of unknown length  
 4980 6965: contig of 1986 bp in length  
 6966 7065: gap of unknown length  
 7066 9365: gap of 2301 bp in length  
 9367 9465: gap of unknown length  
 9467 11957: contig of 2491 bp in length  
 11958 12057: gap of unknown length  
 12058 14208: contig of 2151 bp in length  
 14209 14308: gap of unknown length  
 14309 17022: contig of 2714 bp in length  
 17123 20061: contig of 2939 bp in length  
 20062 20161: gap of unknown length  
 20162 24025: gap of 3864 bp in length  
 24026 24125: gap of unknown length  
 24126 26718: contig of 2593 bp in length  
 26719 26818: gap of unknown length  
 26819 28937: contig of 2118 bp in length  
 28938 29036: gap of unknown length  
 29037 32791: contig of 3755 bp in length  
 32792 32891: gap of unknown length  
 32892 38895: contig of 6004 bp in length  
 38896 38995: gap of unknown length  
 38996 44305: gap of 5311 bp in length  
 44307 44406: gap of unknown length  
 44407 51589: contig of 7183 bp in length  
 51590 51689: gap of unknown length  
 51690 58591: contig of 6902 bp in length  
 58592 58691: gap of unknown length  
 58692 63442: contig of 4751 bp in length  
 63443 63542: gap of unknown length  
 63543 70554: contig of 7012 bp in length  
 70555 70654: gap of unknown length  
 70655 77794: gap of 7140 bp in length  
 77795 77894: gap of unknown length  
 77895 85841: contig of 7947 bp in length  
 85842 85941: gap of unknown length  
 85942 93227: contig of 7286 bp in length  
 93228 93327: gap of unknown length  
 93328 100871: contig of 7544 bp in length  
 100872 100971: gap of unknown length  
 100972 112141: contig of 11170 bp in length  
 112142 112241: gap of unknown length  
 112242 126358: contig of 14117 bp in length  
 126359 126458: gap of unknown length  
 126459 141767: contig of 15309 bp in length  
 141768 141867: gap of unknown length  
 141868 155927: contig of 14060 bp in length  
 155928 156027: gap of unknown length  
 156028 180048: contig of 24021 bp in length







[illegible]

variation /replace="c"  
4061 /gene="IL24"  
/frequency="0.01"  
/replace="a"  
4069 /gene="IL24"  
/frequency="0.36"  
/replace="c"  
4082 /gene="IL24"  
/frequency="0.02"  
/replace="c"  
4095.4167 /rpt\_family="SINE"  
/rpt\_type="dispersed"  
4243 /gene="IL24"  
/frequency="0.37"  
/replace="a"  
4254 /gene="IL24"  
/frequency="0.01"  
/replace="g"  
4332 /gene="IL24"  
/frequency="0.64"  
/replace="a"  
4585 /gene="IL24"  
/frequency="0.17"  
/replace="a"  
4602 /gene="IL24"  
/frequency="0.38"  
/replace="c"  
4943 /gene="IL24"  
/frequency="0.02"  
/replace="a"  
5020 /gene="IL24"  
/frequency="0.01"  
/replace="a"  
5389 /gene="IL24"  
/frequency="0.01"  
/replace="a"  
5392 /gene="IL24"  
/frequency="0.01"  
/replace="a"  
5519 /gene="IL24"  
/frequency="0.01"  
/replace="c"  
5637.5821 /gene="IL24"  
/frequency="0.01"  
/replace="c"  
6016 /rpt\_family="SINE"  
/rpt\_type="dispersed"  
/gene="IL24"  
/frequency="0.35"  
/replace="c"  
6020 /gene="IL24"  
/frequency="0.02"  
/replace="g"  
6037 /gene="IL24"  
/frequency="0.01"  
/replace="c"  
6115.6280 /rpt\_family="SINE"

variation /rpt\_type="dispersed"  
6240 /gene="IL24"  
/frequency="0.01"  
/replace="a"  
6282 /gene="IL24"  
/frequency="0.37"

Query Match 81.98; Score 1872.2; DB 9; Length 11114;  
Best Local Similarity 98.68; Pred. No. 0;  
Matches 1916; Conservative 0; Mismatches 8; Indels 19; Gaps 2;

QY 363 gtaactgaagattcagccgcccctggggtcatttccctcctaagctcgttaagc 422  
Db 1 gfaactgaagatttcagcgcctcctggggtcatttccctcctaagctcgttaagc 60  
QY 423 tgtctcctgcctcccaactaactatctacgaactcgaactcgaactcctctctc 482  
Db 61 tgtctcctgcctcccaactaactatctacgaactcgaactcgaactcctctc 120  
QY 483 lcaactgcctcctcgaactatctacgaactcgaactcgaactcgaactcctctc 542  
Db 121 tcactgcctcctcgaactatctacgaactcgaactcgaactcgaactcctctc 180  
QY 543 cctaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 602  
Db 181 cctaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 240  
QY 603 ctactcctggaaggtgaagactcgaactcgaactcgaactcgaactcgaactcctctc 662  
Db 241 ctactcctggaaggtgaagactcgaactcgaactcgaactcgaactcgaactcctctc 300  
QY 663 taagcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 722  
Db 301 taagcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 360  
QY 723 taagcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 782  
Db 361 taagcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 420  
QY 783 gtcgccaaggtggttctcactcactcctcctcctcctcctcctcctcctcctcctc 842  
Db 421 gtcgccaaggtggttctcactcactcctcctcctcctcctcctcctcctcctcctc 480  
QY 843 gcaagtaactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 902  
Db 481 gcaagtaactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 540  
QY 903 tgagcaaacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 962  
Db 541 tgagcaaacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 600  
QY 963 ggtcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 1022  
Db 601 ggtcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 660  
QY 1023 gtcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 1082  
Db 661 gtcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 720  
QY 1083 ataagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1142  
Db 721 ataagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 780  
QY 1143 tcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 1202  
Db 781 tcgaagcagactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 840  
QY 1203 agttggccaatccctcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 1262  
Db 841 agttggccaatccctcgaactcgaactcgaactcgaactcgaactcgaactcgaactcctctc 900

```

QY 1263 gaatgtgattaaacagagatlaaagaagcatcacatctccaggaagcgttagacat 1322
DB 901 GAATGTGATTAAACAGAGATTAAAGACATCACATCTTCCAGAGAGGGCTAGCAT 960
QY 1323 gcaagagctcttaactgaactcaat-agaatcaatctgttcaatggaagatcttgaca 1381
DB 961 GCCAGGCTCTTAACTGACATCATATAGATTCATCTTTCTATGAGAGATCTTGACAGA 1020
QY 1382 caagaattcccaaacctagaagctgagagagctgtcttggaagaaacgtctgctga 1441
DB 1021 CAAGAATCTCTCAACCTAGAGCTGAGAGAGCTGCTTTGGGAAACATGCTGCTTGA 1080
QY 1442 tgcctcaatgagacatgtagaattcaagagctgagtgctgtcagctggaagatggg 1501
DB 1081 TGCCCTGACGGGACATGATGATCTTAAGACTAGAGCTGCTCTCTACCTGGAATGAGG 1140
QY 1502 tcaagagctctgagctgagcagctgagctgagacacacagagctgagctgagctgag 1561
DB 1141 TCAGAGCTCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1200
QY 1562 gttcagagacacagagcagaagctagaccacagagacacagagcctccagcttctg 1621
DB 1201 GTTCAGAGCCACACAGGCGACAGCTAGCCACAGAGACAGAGCTGCTGCTTCTTG 1260
QY 1622 gctgtcagagctagacattctctctctgctgcttccagaaactagaagagagcaca 1681
DB 1261 GCTGTGCTAGCTGATTCCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1682 ttttggtgagtgattataaccttagagacatggtgagctgagctgagctgagctgag 1741
DB 1321 TTTTGGTGGATGATTAACCTTAGGAGACATGCTGAGCTGAGCTGAGCTGAGCTGAG 1380
QY 1742 tcaactccctgagcctgagagatlaaaggagagagctgagctgagctgagctgagct 1801
DB 1381 TCAACTCTCTGAGCTGAGATTAAGAGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1440
QY 1802 gactacgtctccagagctagatagagacatcatattgaatttttgaagagagagagact 1861
DB 1441 GACTACGTCTCCAGCATGATGAGGAGATCATATTGATTAATTGATGAGAGAGAGACT 1500
QY 1862 ggtatgctgacagaglaaataatgattacatgctgacagagcttaccagagcacttg 1921
DB 1501 GGTGTATGCTGACAGATTAATGATTAATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1922 tttgttttgccttggttattacactt-----gggagcaattttc 1963
DB 1561 TGTGTTTGTGCTCTGTTTATTACACTTGCCCTTGGGGGAGGTGAGCAAAATTTTGA 1620
QY 1964 aaaaattatacatgagagagctgacagcagagagagctaaagagactgcccctgcccaca 2023
DB 1621 AAAATTATATGACAGAGAGCTGACAGGAGAGAGAGCTMAAGAGACTGCCCTGCCACA 1680
QY 2024 cagcagatgtagagagctgaactcaaacagagctgacatctcaactcagagagagctgct 2083
DB 1681 CAGCAGAGGAGAGAGCTGAGCTCAAAAGAGAGCTGATCTCACTCAAGGAGGCTGCTTC 1740
QY 2084 cccatcgtctgattgctcctaaagtgatgtagctgagcagaaagaaatctcctaga 2143
DB 1741 CCCATGCTGATGATGCTTAAAGATGATGAGTACAGCAATGAATTAATCTCTAGA 1800
QY 2144 aagatgacaaatttcccttctccacccctccttcttctccacccctcccatcag 2203
DB 1801 AAGATGACAAATTTCCCTTCTCCACCTCCCTCTTCTCTCCACCTCCCTCCCATAG 1860
QY 2204 ccccatatataatgcccacaaatctccacaaagcctgctgctgacaaacttactctg 2263
DB 1861 CCCCATAATATATGCCCAATCTCCACAAAGCTGCTGCTGCAAAACCTTACTCTG 1920
QY 2264 aatgactctcagagctgagagc 2286
DB 1921 AAATGACTCTCAGGCTGGAGC 1943

```

```

RESULT 6
AF276916 7025 bp DNA linear PRI 20-DEC-2000
LOCUS AF276916
DEFINITION Homo sapiens melanoma differentiation association protein 7 gene,
complete cds.
VERSION AF276916
KEYWORDS AF276916.1 GI:11907920
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7025)
AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.
TITLE The human MDA-7 gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7025)
AUTHORS Peat,J., Kube,D., Eskdale,J., Jueliger,S. and Gallagher,G.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-2000) Department of Surgery, University of
Glasgow, Queen Elizabeth Building, Glasgow Royal Infirmary, Glasgow
G3 7ER, Scotland
FEATURES
source
1..7025
/organism="Homo sapiens"
/db_xref="taxon:9606"
join(<1804..1847,3285..3480,4261..4323,5460..5618,
5964..6038,6942..>7025)
/product="melanoma differentiation association protein 7"
join(1804..1847,3285..3480,4261..4323,5460..5618,
5964..6038,6942..7025)
/codon_start=1
/product="melanoma differentiation association protein 7"
/protein_id="AAG41401.1"
/db_xref="GI:11907921"
/translation="NMFOORLOSMTLAPRCPPPLAVASOMQVYVPLCGFTLLMS
QVAGAGQEPHFGPCQVGVYVPOKIMEFAMVAKDMOAOHITSAKLQOEVLQVSD
AESCVYHLLRFLYKATVRYKYNHNETVGRVRLKSPSLANNFVITLSQLOPSENMFM
SINDSHRRPLFRRAFKQDYEAALTALEVEVILLRWKPKKL"
BASE COUNT 1759 a 1668 c 1759 g 1838 t 1 others
ORIGIN
Query Match 59.7%; Score 1365.2; DB 9; Length 7025;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1425; Conservative 0; Mismatches 14; Indels 22; Gaps 3;
QY 848 taccctgtccccaagtcacattcctactccglttccacagcctggaagcaggttggc 907
DB 1 TATCTGTGCCCCACGTCACATTCCTCCCTTCCATGCTGAGCCAGGTTGGGC 60
QY 908 aaactctctgttaagagacagagaaactatttaggcctgttgacatatgctc 967
DB 61 AAACCTTCTCTTAAGAACACAGACAGAACTATTGAGCGCTGTGSCATATAGCT 120
QY 968 cagtcacacactcactcctcctcctgtgacagaaagaaatagcaaatatgctca 1027
DB 121 CAGTCACAATCACTCACTGCTGTGACAGAAACAAATTGCAACATATGTCAA 180
QY 1028 caaacatagtgacccatgaaaacttatattatgtgatacgaagaaactgaaaataat 1087
DB 181 CAATATATGTGACCCCATGAAACCTTATTATTATTTGATATGGAACGAAACCTGAATATAT 240
QY 1088 gctcttcttctgatttlltcccaatcataaanaacglaaanaactacacttaagtcga 1147
DB 241 GCTCTTCTTTGATTTTTCCTCCATCATTAATAAAGCAATAGCTTACCTGACCA 300
QY 1148 aggttaagcactctcagcttagagagagcagctgagattgctgctgacataagctg 1207
DB 301 AGGTTAAGCATTCTCACTTACAGTGGAGGCTGATTTGGCTTGTGACCTACAGTTG 360
QY 1208 gccaatccctgattcccaaaatgattctcctcagagatgctgagcaaatatggaagt 1267
DB 1208 gccaatccctgattcccaaaatgattctcctcagagatgctgagcaaatatggaagt 1267

```

Dd	361	GGCAATTCCTGATTGATCCAAAATGATATTTCCTCAAGGATGTGGGCAAAATACATTAATGGCAAAAT	420
Qy	1268	gctggagatctaaagaagatlaagaagagcttgagaaattccagaagagggctagacatgacag	1327
Dd	421	GCTGGATTAAACAGAGTTAAAGAACATCAAGACATTTTCCAGAGCGGCTAGACATGCCAG	480
Qy	1328	ggtctctcaatgacatcatt-ggattcaatcgttctcatgtaggaaccttgcaagaaga	1386
Dd	481	GGCTCTCTAACTGACCTCATATTAGAAATTCATCGTTTCATGAGAGATCTTGCAAGACAAGA	540
Qy	1387	atcccccanaaccttagagctctgagagcttgctcttgaggaaacaacgctcgtctatgagcc	1446
Dd	541	ATTTCCTAAACCTAAGACTAGCTAGAGCTGAGCTGTGGGAAACACAGCTCTCTGTGATGCC	600
Qy	1447	tcacttgagacatgtagaatactagaagctgagatgacctgtcagctcagcttgagaataagatcaga	1506
Dd	601	TCACCTGGCACATGATGAATCTAGAGCTGATGCTCTGCTACGTGGATAGATAGGCTGAGA	660
Qy	1507	gctcttgactgccttgagcagctctgacaacatcagcgtctcgtgtccctgagtggttca	1566
Dd	661	GGCTTGACCTGCCCTGGAGAGCTTGACATACATACGCTCTGTGTCCCTCGAATGGTTTCA	720
Qy	1567	gaagcaacaaggccaaagacttagccccaaccaagacacagagccctccagctctctggagctg	1626
Dd	721	GAGCCACACAGGCGCAAGACTAGCCCAACACAGACACAGGCTCCACACTTTCGGGCTTG	780
Qy	1627	tcgaatgagatacttccattcttccctggtcttcagaagccaaagagagagacatttgg	1686
Dd	781	TCGATCGTACATTTCTCTATTCTCTCTGTTTCCAGAACTTAAGAGAGACGACATTTTG	840
Qy	1687	gttgagtgatcttaaccttaggacatgagtagctgacatgatacagaacaactctccaac	1746
Dd	841	GTTGACATGTATTAACCCCTAGGACACATGGGATAGCTGATGTCAGAAACATCTCAAC	900
Qy	1747	ctctctggacctgaatgatactaaagagagtaacttcaagatlaattctctcgtct--ggg	1803
Dd	901	TTCTCTGGCCCTATGATTAAGAGAGAGCTTACAGTTATTCTTCGCTATTAACGA	960
Qy	1804	ctctctgtcccaagacagaaatagggaatcatattgaattattttgagagagagagactgg	1863
Dd	961	CTACCTGCCACAGATGAATAGGAGCATCATTTTGAATATTTTTACAGAGAGAGAGCTGG	1020
Qy	1864	tgtatgctgacagtaataatgatacthaatgatacagtgatacagattcaagaacactctgtg	1922
Dd	1021	TGATATCTGCACAGTAATAATGATATTTCATGATGTAAGAGATTACCAACACTCTGTGG	1080
Qy	1924	tgttttttgccttcttcttaatacactl-----gggaaanaattctlaa	1965
Dd	1081	TTGTTTTTCCCTCTTATTAATACCTATGCCCTTGGGGAGAGTGGGCAAAATTTTATAA	1140
Qy	1966	aattctaaatgacaaagactcgaagcagaagagactaaagaagcttgacccctgcacaca	2022
Dd	1141	AATTTATCATCTCAGAGACTGACAGCCAGACAGAAAGTAAGAACTTGCCCTGGCCCAKCA	1200
Qy	2026	gccagtgatagagcttgaactcaaaccccaagctcatctcaactcaagggcgtgcttccc	2085
Dd	1201	GGCAGTGGTGAAGCCTGTGAACCAACCCAGGCGCTATCTCACTCAAGGCGCTCTTTCC	1260
Qy	2086	cacagctcatgtctcttaagtgatggatgactagagaagaagtaattctcttagaaga	2145
Dd	1261	CATCGCTGAATGGCCCTTAAAGTATGATGGGACTAGCGAATGAAGTAAATCTCTAGAAA	1320
Qy	2146	gcatgacaaattccctctctccaactctcctctttctccaaacctcccccaatcagcc	2205
Dd	1321	GCATGACAAATTTCCCTTTCTCCACCTCCCTCTTTTCTCCACCCCTCCCATAGGC	1380
Qy	2206	cccaataataatgccaaatctccacaagaagctgcttgccgcgcaaaccttatacttga	2265
Dd	1381	CCCATATATATGCCCAATCTCCACAAAGCTGTGCTGCTGCAAACTTTACTTTGTGA	1440
Qy	2266	atgactccaagagctgggaag 2286	
Dd	1441	ATTGATTTCCAGGCTGGGACG 1461	

RESULT	7
LOCUS	HSZA243947
DEFINITION	Homo sapiens VCX-A gene for variably charged X chromosome protein.
ACCESSION	AJ243947
VERSION	AJ243947.1 GI:5912555
KEYWORDS	VCX-A gene.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 29034) Li,X.M., Yen,P.H. and Shapiro,L.J. Characterization of a low copy repetitive element 523z involved in the generation of frequent deletions of the distal short arm of the human X chromosome Nucleic Acids Res. 20 (5), 1117-1122 (1992)
JOURNAL	92195814
MEDLINE	2 (bases 1 to 29034)
REFERENCE	Fukami,M., Kirsch,S., Schiller,S., Richter,A., Benes,V., Franco,B., Fukaya,K., Rao,E., Merker,S., Niesler,B., Ballabio,A., Ansoorge,W., Ogata,T. and Rapin,I.G.A. A member of a gene family on Xp22.3, VCX-A, is deleted in patients with x-linked nonspecific mental retardation Am. J. Hum. Genet. 67 (3), 563-573 (2000) 20365828
TITLE	3 (bases 1 to 29034) Richter,A. Direct Submission Submitted (10-AUG-1999) Richter A., Biochemical Instrumentation, EMBL, Meyerhoistrasse 1, Heidelberg, 69117, GERMANY Location/Qualifiers 1..29034 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="Xp22.3" /map="between 40Bp and DXS1139" /clone="TLCN110" /sex="male"
FEATURES	/cell_line="lymphoblastoid" /clone_id="Lawrence Livermore X-chromosome specific library" note="Variably charged, X chromosome mRNA on CMT-S232A" 10574..10732 /gene="VCX-A" number=1 10574..11947 /gene="VCX-A" 10733..11046 /gene="VCX-A" number=1 11047..11294 /gene="VCX-A" number=2 join(11195..11294,11487..11947) /gene="VCX-A" codon_start=1 product="VCX-A protein" protein_id="CA56197.1" /db_xref="GI:5912556" translation="MSPKPRASGPPAKATEAGKRSSQSPSPDKTKTTVYAKKGA VRGRGRKKGATMAAVTPAPEAESGPAPRGSDPSQLPHHELPRPEPSECTQID PLDSQLEELPELQSVEEPLDSQVEEPPLDSQVEEPLDSQVEEPLDSQVEEPLD EPIPSQSEEEPLDSQSEMEELPSV" 11295..11486 /gene="VCX-A" number=2 11487..11947 /gene="VCX-A" number=3 11697..11936
repeat_region	





Oy	1032		catactgacccaccatbaaacatttttattatcatgaacagcaaacccaataaatgcctc	1091
Db	44870	TAGGGGACACTCTCCAAATGAACAACATTGTTCACAGTCTCAAGAAGTTGCATTATATATA	44811	
Oy	1092	tctct	1095	
Db	44810	TTTT	44807	
RESULT	10			
LOCUS	AL139275	193126 bp	-DNA	linear PRI 04-SEP-2001
DEFINITION	Human DNA sequence from clone RP11-552E20 on chromosome			
ACCESSION	AL139275			
VERSION	ALI39275.30	GI:15485119		
KEYWORDS	HTG.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eskayrova; Merazou; Choradala; Cranialta; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 193126)			
JOURNAL	Smith,M.			
COMMENT	Direct Submission Submitted (04-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Requests: clonerequests@sanger.ac.uk On Sep 6, 2001 this sequence version replaced gi:14800199. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <a href="http://www.sanger.ac.uk/projects/C_elegans/wormpep">http://www.sanger.ac.uk/projects/C_elegans/wormpep</a> This sequence was generated from part of bacterial clone contigs of human Chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/HGP/Chr6">http://www.sanger.ac.uk/HGP/Chr6</a> RP11-552E20 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a> VECTOR: pBACe3.6			
FEATURES				
source				
	1..193126	/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/chromosome="6"		
		/map_pos=3-21.2"		
		/clone="RP11-552E20"		
		/clone_lib="RPCI-11.2"		
	12..455	/note="match: GSS: Em:A0807014"		
	4253..4780	/note="match: GSS: Em:B71338"		
	4256..4678	/note="match: GSS: Em:AA437688"		
		/complement(8999..9366)		



misc_feature	/note="match: SMS: Em: G63453" 9829. .10237
misc_feature	/note="match: GSS: Em: A0030932" complement(10116. .106117)
misc_feature	/note="match: SMS: Em: G66076 match: GSS: Em: A0318725"
misc_feature	13203. .13321
misc_feature	/note="match: GSS: Em: A0792274" 13238. .13428
misc_feature	/note="match: GSS: Em: A0042370" complement(13332. .13469)
misc_feature	/note="match: GSS: Em: A0504765" 15417. .15794
misc_feature	/note="match: GSS: Em: A0138303" 15426. .15971
misc_feature	/note="match: GSS: Em: A0384582" 15435. .16067
misc_feature	/note="match: GSS: Em: A0198235" 20235. .20738
misc_feature	/note="match: GSS: Em: A0480601" 20275. .20620
misc_feature	/note="match: GSS: Em: A0097268" 33115. .33591
misc_feature	/note="match: GSS: Em: A0613320" complement(33079. .33571)
misc_feature	/note="match: GSS: Em: A0692857" complement(38680. .39467)
misc_feature	/note="match: GSS: Em: A0092878" 41966. .42582
misc_feature	/note="match: GSS: Em: A0308209" 41976. .42381
misc_feature	/note="match: GSS: Em: A0124355" 46566. .46984
misc_feature	/note="match: GSS: Em: A0882156" 47399. .47848
misc_feature	/note="match: GSS: Em: A0811844" 64782. .65002
misc_feature	/note="match: GSS: Em: A0772609" 64783. .65002
misc_feature	/note="match: GSS: Em: A0683445" complement(64833. .64883)
misc_feature	/note="match: SMS: Em: G08189" complement(64864. .64983)
misc_feature	/note="match: GSS: Em: A1328221" complement(64864. .64975)
misc_feature	/note="match: GSS: Em: A1324222" complement(64888. .64983)
misc_feature	/note="match: GSS: Em: A0332064" 64898. .65015
misc_feature	/note="match: SMS: Em: G09742" complement(64898. .64901)
misc_feature	/note="match: GSS: Em: A257084" complement(64899. .64989)
misc_feature	/note="match: GSS: Em: A0674454" complement(64900. .64989)
misc_feature	/note="match: GSS: Em: A237938" 64903. .65013
misc_feature	/note="match: GSS: Em: A191898" 64903. .64991
misc_feature	/note="match: GSS: Em: A1309316 Em: A1316981 Em: A1346021" 64903. .64983
misc_feature	/note="match: GSS: Em: A1224697" 64903. .64993
misc_feature	/note="match: GSS: Em: A1284370" complement(64903. .64983)
misc_feature	/note="match: SMS: Em: G10109" 64905. .65002
misc_feature	/note="match: GSS: Em: A0284087" complement(74424. .74751)
misc_feature	/note="match: GSS: Em: A025184" complement(77134. .77547)
misc_feature	/note="match: GSS: Em: A0192713" 80797. .81303

	/note="match": GSS: Em:A0606590"
misc_feature	/complement(80886...81417)
	/note="match": GSS: Em:A0602297"
misc_feature	/complement(91202...91736)
	/note="match": GSS: Em:A0065130"
misc_feature	/complement(102120...102521)
	/note="match": GSS: Em:A0177699"
misc_feature	107123...107721
	/note="match": GSS: Em:A0524997"
misc_feature	/complement(115935...116552)
	/note="match": GSS: Em:A0080194"
misc_feature	/complement(116657...116478)
	/note="match": GSS: Em:A0031004"
misc_feature	/complement(116193...116553)
	/note="match": GSS: Em:B32716"
misc_feature	/complement(128470...128806)
	/note="match": GSS: Em:B36077"
misc_feature	/complement(128811...129176)
	/note="match": STS: Em:HS164XP4"
misc_feature	139541...139825
	/note="match": GSS: Em:A0278260"
misc_feature	165778...165946
	/note="match": STS: Em:G26303"
misc_feature	171874...172092
	/note="Single clone region. Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."
misc_feature	/complement(175607...176503)
	/note="match": GSS: Em:A0882642"
misc_feature	176202...176743
	/note="match": GSS: Em:A0374550"
misc_feature	/complement(177911...178307)
	/note="match": GSS: Em:A0131194"
misc_feature	/complement(177997...178301)
	/note="match": GSS: Em:A0148743"
misc_feature	178325...178821
	/note="match": GSS: Em:A0420001"
misc_feature	/complement(179120...179217)
	/note="match": GSS: Em:A0545214"
BASE COUNT	54062 a 43323 c 43111 g 52630 t
ORIGIN	

Query Match	4.0%	Score 90.8	DB 9	Length 19126
Best Local Similarity	55.7%		Pred. No. 3.5e-14	
Matches 194	Conservative	0	Mismatches 122	Indels 2
			Gaps	1
QY 901	gtcggcaaacctccctcgtlaaagaacacgaagaacatatttagctctcgtgacct	960		
Db 125379	GGTCAGCAAAACATTTTCTGTGAAGAACCAGATAGCATATTTTATGGCTTTGGGGCCAT	125438		
QY 961	atgcttcagatgacacactcactcctcctcctcgtctgacgaagaagaattagacaacat	1020		
Db 125439	ATTAATCTCTGTACACACTGCTACCTCCGCCATTCGACTGTGAAAGCAGCTGTAGACAAA	125498		
QY 1021	atgtcaacaacatctgacccacatgaaacattattattatgatacgaagaacctga	1080		
Db 125489	ACGTAATATGAACGGGATGCTGTGCCACATGAATACTTTGGGACCATGAATAATTAA	125558		
QY 1081	aaataaagctctctctcttctgattctttccccaatctaaanaaagtaanaacacacctta	1140		
Db 125559	TTTTTCATGCTGCACAAAGATTTTCTTTTGTGATTTTCCACACCAATTAATAATGTAAAAAC	125618		
QY 1141	ggttcgaagaagttlaagcaatcttcagcttagtagcgtgacgcgtgatttgctctgtgacct	1200		
Db 125619	CTTTCCTTAACCTTAGGGGTAGAACAAATACAG--GCATCGGGCTGGATTTGGCTCATGCGCC	125676		
QY 1201	acagcttgagcaatccctcctgattcccaaaatgatactcttagagtagctg	1248		
Db 125677	ATTGTGTCGCCACCTGATTGACATCATCCCATCTACAAATGTGG	125724		



LOCUS	AC025298	17957 bp	DNA	linear	HTG 03-JUN-2000
DEFINITION	Homo sapiens clone Rpl1-2N21, WORKING DRAFT SEQUENCE, 5 unordered pieces.				
ACCESSION	AC025298				
VERSION	AC025298.3	GI:8225254			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.				
TITLE	Homo sapiens, clone Rpl1-2N21				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 177957)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barre, N., Bastien, V., Bede, F., Boguski, L., Bouhassira, B., Brown, A., Burkett, G., Caspi, A., Castle, A., Chao, P., Chao, P., Collins, S., Collins, S., Cooke, P., Dearfield, K., Dewar, K., Diaz, J.S., Dode, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargay, S., Ginde, S., Goyette, M., Graham, L., Grand, P., Grant, N., Grant, G., Hago, B., Heaford, A., Horton, L., Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J., Lacroque, K., Lamazeres, R., Landers, T., Lehotzky, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McKernan, P., McGuck, A., McKernan, K., McPheters, R., Melid, J., Menes, L., Mihova, T., Miranda, C., Miya, J., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olliver, T.M., Oliver, J., Peterson, K., Piatre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Sever, P., Spencer, B., Stange, H., Stojanovic, N., Subramanian, A., Talmas, J., Tassilev, S., Theodor, J., Tirelli, A., Travers, M., Triggilio, J., Vassilev, H., Veld, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.				
TITLE	Direct Submissions				
JOURNAL	Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
COMMENT	On Jun 3, 2000 this sequence version replaced gi:7670170. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html				
-----					
Center: Whitehead Institute/ MIT Center for Genome Research					
Center code: W18R					
Web site: http://www-seq.wi.mit.edu					
Contact: sequence.submissions@genome.wi.mit.edu					
-----					
Project Information					
Center project name: L2680					
Center clone name: 2.N.21					
-----					
Summary Statistics					
Sequencing vector: M13; M71815; 100% of reads					
Chemistry: Dye-terminator Big Dye; 100% of reads					
Assembly program: Phrap; version 0.960731					
Consensus quality: 174723 bases at least Q40					
Consensus quality: 176598 bases at least Q30					
Consensus quality: 177197 bases at least Q20					
Insert size: 183000; agarose-fp					
Insert size: 177557; sum-of-contigs					
Quality coverage: 6.7 in Q20 bases; agarose-fp					
Quality coverage: 6.8 in Q20 bases; sum-of-contigs					
-----					
NOTE: This is a 'working draft' sequence. It currently					
* consists of 5 contigs. The true order of the pieces					
* is not known and their order in this sequence record is					
* arbitrary. Gaps between the contigs are represented as					
* runs of N, but the exact sizes of the gaps are unknown.					
* This record will be updated with the finished sequence					
* as soon as it is available and the accession number will					
* be preserved.					
1					
1828: contig of 1828 bp in length					





VERSION AC011060.5 GI:7229784  
 KEYWORDS HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 181009)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens, clone RP11-11616  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 181009)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barua,N., Beckert,L., Boguslavsky,L., Bouckgatter,B.,  
 Brown,A., Castle,A., Colangelo,R., Collins,S., Collymore,A.,  
 Cooke,P., Dearliano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gardina,S., Grant,C., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kamp,L., Karatas,A., Klein,J.,  
 Lebecky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tirrell,A., Vasilev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome  
 Research, 330 Charles Street, Cambridge, MA 02141, USA  
 COMMENT On Mar 12, 2000 this sequence version replaced g1:6984405.  
 All repeats were identified using RepeatMasker:  
 hmtc, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

Project Information  
 Center project name: L3156  
 Center clone name: 11\_G\_16

Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 163981 bases at least Q40  
 Consensus quality: 172073 bases at least Q30  
 Consensus quality: 175479 bases at least Q20  
 Insert size: 155000; agarose-fp  
 Insert size: 178609; sum-of-contigs  
 Quality coverage: 5.1 in Q20 bases; agarose-fp  
 Quality coverage: 4.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
 consists of 25 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1504: contig of 1504 bp in length  
 1505 1604: gap of 100 bp  
 1605 2791: contig of 1187 bp in length  
 2792 2891: gap of 100 bp  
 2892 3909: contig of 1018 bp in length  
 3910 4009: gap of 100 bp  
 4010 5556: contig of 1547 bp in length  
 5557 5656: gap of 100 bp  
 5657 7482: contig of 1826 bp in length  
 7483 7582: gap of 100 bp  
 7583 8926: contig of 1344 bp in length  
 8927 9026: gap of 100 bp  
 9027 10470: contig of 1444 bp in length

## FEATURES

## SOURCE

misc_feature	10471	10570:	gap of 100 bp
misc_feature	10571	12199:	contig of 1629 bp in length
misc_feature	12200	12299:	gap of 100 bp
misc_feature	12300	14087:	contig of 1768 bp in length
misc_feature	14088	14187:	gap of 100 bp
misc_feature	14188	15707:	contig of 1520 bp in length
misc_feature	15708	15807:	gap of 100 bp
misc_feature	15808	17945:	contig of 2138 bp in length
misc_feature	17946	18045:	gap of 100 bp
misc_feature	18046	20017:	contig of 1972 bp in length
misc_feature	20018	20117:	gap of 100 bp
misc_feature	20118	22115:	contig of 1998 bp in length
misc_feature	22116	22215:	gap of 100 bp
misc_feature	22216	24556:	contig of 2341 bp in length
misc_feature	24557	24656:	gap of 100 bp
misc_feature	24657	27568:	contig of 2912 bp in length
misc_feature	27569	27668:	gap of 100 bp
misc_feature	27669	32527:	contig of 4859 bp in length
misc_feature	32528	32627:	gap of 100 bp
misc_feature	32628	37892:	contig of 5265 bp in length
misc_feature	37893	37992:	gap of 100 bp
misc_feature	37993	41265:	contig of 3273 bp in length
misc_feature	41266	41365:	gap of 100 bp
misc_feature	41366	52400:	contig of 11035 bp in length
misc_feature	52401	52500:	gap of 100 bp
misc_feature	52501	63506:	contig of 11006 bp in length
misc_feature	63507	63606:	gap of 100 bp
misc_feature	63607	79362:	contig of 15756 bp in length
misc_feature	79363	79462:	gap of 100 bp
misc_feature	79463	97023:	contig of 17561 bp in length
misc_feature	97024	97123:	gap of 100 bp
misc_feature	97124	124426:	contig of 27303 bp in length
misc_feature	124427	124526:	gap of 100 bp
misc_feature	124527	149337:	contig of 24811 bp in length
misc_feature	149338	149437:	gap of 100 bp
misc_feature	149438	181009:	contig of 31572 bp in length.

Location/Qualifiers  
 1. 181009  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-11616"  
 /clone\_id="RP11-11616"  
 1. 1504  
 /note="assembly-fragment"  
 1605. 2791  
 /note="assembly-fragment"  
 2892. 3909  
 /note="assembly-fragment"  
 4010. 5556  
 /note="assembly-fragment"  
 5557. 7482  
 /note="assembly-fragment"  
 7483. 8926  
 /note="assembly-fragment"  
 8927. 10470  
 /note="assembly-fragment"  
 10471. 10570  
 /note="assembly-fragment"  
 10571. 12199  
 /note="assembly-fragment"  
 12200. 14087  
 /note="assembly-fragment"  
 14088. 14187  
 /note="assembly-fragment"  
 14188. 15707  
 /note="assembly-fragment"  
 15708. 15807  
 /note="assembly-fragment"  
 15808. 17945  
 /note="assembly-fragment"  
 17946. 18045  
 /note="assembly-fragment"  
 18046. 20017  
 /note="assembly-fragment"  
 20018. 20117  
 /note="assembly-fragment"  
 20118. 22115  
 /note="assembly-fragment"  
 22116. 24556  
 /note="assembly-fragment"  
 24557. 24656  
 /note="assembly-fragment"  
 24657. 27568  
 /note="assembly-fragment"  
 27569. 27668  
 /note="assembly-fragment"  
 27669. 32527  
 /note="assembly-fragment"

```
misc_feature /note="assembly_fragment"
32628..37892
/note="assembly_fragment"
misc_feature 37993..41265
/note="assembly_fragment"
clone_end:SP6
misc_feature 41365..52400
/note="assembly_fragment"
52501..63506
/note="assembly_fragment"
63607..79362
/note="assembly_fragment"
clone_end:77
vector_side:right"
misc_feature 79463..97023
/note="assembly_fragment"
79124..124426
/note="assembly_fragment"
124527..149337
/note="assembly_fragment"
149438..181009
/note="assembly_fragment"
BASE COUNT 57610 a 36532 c 33173 g 51292 t 2402 others
ORIGIN
```

Query Match 3.9%: Score 89.2; DB 2; Length 181009;

Best Local Similarity 64.6%: Pred. No. 9.9e-14; Matches 133; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy 891 ctgagccaggttggcgaactctcctgtaagaacagacaggaactatttagctc 950
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151319 CTGACAGAGAGCGCAGCAACTTCTCTGTAAGAGACAGATGATTAATTTAGGTT 151378

Qy 951 tctgtgcataatggtctcagtcacaactactcatctcctctgtgacgcgaagaat 1010
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151379 TGTGGGCGCATATGCTCTCTGTCACAACTCACTGCTGCTTCTAGGGCAAGCAGT 151438

Qy 1011 tagaacaataatgtaacacaacataatgtgaccccatgaacattatattatgagtaac 1070
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151439 TATAGAAATATACATATATGATGACATGCTGTGTTCCCATAAAGATTATTATGACAC 151498

Qy 1071 ggaacctgaaataatgtcttctt 1096
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151499 TGAATTTTGAATTCATGTAATTTT 151524
```

```
RESULT 15
AL161628 183483 bp DNA linear PRI 16-JAN-2001
LOCUS Human DNA sequence from clone RP11-31K16 on chromosome 9. Contains
DEFINITION a snRNA binding domain pseudogene, the ELAVL2 gene for ELAV
(embyronic lethal, abnormal vision, Drosophila)-like 2, ESTs, STS,
GSSs and a Cpg island, complete sequence.
ACCESSION AL161628
VERSION AL161628.9 GI:10129841
KEYWORDS HNG: Cpg island; ELAVL2; snRNA binding domain.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 183483)
AUTHORS Ramsay, H.
JOURNAL Direct Submission
Submitted (20-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
UK
requests: clonerequests@sanger.ac.uk
On Sep 14, 2000 this sequence version replaced gi:10039694.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

This sequence is the entire insert of clone RP11-31K16. The true left end of clone RP11-32IL2 is at 139671 in this sequence. The true right end of clone RP11-315114 is at 73045 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-31K16 is from the library RPc1-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

#### FEATURES

##### source

```
1..183483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-31K16"
/clone_id="RPc1-11.1"
1..224
/note="match: GSS: Em:A0040596"
1..456
/note="match: GSS: Em:A0080076"
24..653
/note="match: GSS: Em:A0045921"
415..471
/note="Alu10 repeat: matches 60..118 of consensus"
472..763
/note="Alu1 repeat: matches 10..288 of consensus"
764..954
/note="Alu10 repeat: matches 118..297 of consensus"
1096..1173
/note="MIR10 repeat: matches 103..187 of consensus"
2555..4081
/note="L1MC2 repeat: matches 4721..6314 of consensus"
4184..4223
/note="MIR11 repeat: matches 238..273 of consensus"
4424..4479
/note="MIR11 repeat: matches 343..398 of consensus"
5387..5466
/note="MER58 repeat: matches 2214..2294 of consensus"
5639..5786
/note="L1M3A repeat: matches 6017..6156 of consensus"
6111..6324
/note="MER20 repeat: matches 1..218 of consensus"
6330..6457
/note="MER5A repeat: matches 2..120 of consensus"
6840..7315
/note="L2 repeat: matches 2252..2736 of consensus"
7345..7434
/note="MIR repeat: matches 2..89 of consensus"
7437..7741
/note="AluX repeat: matches 1..299 of consensus"
7808..7989
/note="MER58 repeat: matches 28..2412 of consensus"
10366..10649
/note="MER58B repeat: matches 46..341 of consensus"
11076..11221
/note="MIR2 repeat: matches 1..147 of consensus"
11222..11661
/note="MIR2CB repeat: matches 1..504 of consensus"
11664..11992
```

```
repeat_region /note="MLT2 repeat: matches 145. .510 of consensus"  
12210. .12467  
misc_feature /note="12 repeat: matches 2482. .2750 of consensus"  
complement(12686. .13167)  
misc_feature /note="match: GSS: Em:AQ549073"  
complement(13792. .14266)  
misc_feature /note="match: GSS: Em:AQ594257"  
complement(14438. .14735)  
repeat_region /note="match: GSS: Em:AQ128238"  
14756. .14785  
repeat_region /note="15 copies 2 mer aa 100% conserved"  
16212. .16363  
repeat_region /note="FRAM repeat: matches 4. .145 of consensus"  
16364. .16385  
misc_feature /note="11 copies 2 mer tt 100% conserved"  
16462. .16951  
repeat_region /note="match: GSS: Em:AQ631149"  
16908. .17128  
misc_feature /note="MIR repeat: matches 34. .261 of consensus"  
complement(17116. .17552)  
misc_feature /note="match: GSS: Em:AQ133240"  
17546. .18050  
misc_feature /note="match: GSS: Em:AQ777145"  
17660. .17960  
repeat_region /note="match: GSS: Em:AQ094704"  
19777. .19993  
misc_feature /note="MER58A repeat: matches 2. .224 of consensus"  
20817. .21173  
misc_feature /note="match: GSS: Em:AQ376698"  
complement(20830. .20939)  
misc_feature /note="match: GSS: Em:AQ229933"  
complement(20830. .21133)  
misc_feature /note="match: GSS: Em:AQ235822"  
complement(20830. .21137)  
repeat_region /note="match: GSS: Em:B65379"  
21100. .21154  
repeat_region /note="11 copies 5 mer aaac 78% conserved"  
21183. .21410  
repeat_region /note="MER30 repeat: matches 1. .230 of consensus"  
21647. .21773  
misc_feature /note="MIR repeat: matches 119. .246 of consensus"  
complement(23139. .23500)  
repeat_region /note="match: GSS: Em:AQ210496"  
23363. .23601  
misc_feature /note="MIR repeat: matches 2. .256 of consensus"  
complement(24372. .24810)  
repeat_region /note="match: GSS: Em:AQ178605"  
26871. .27202  
repeat_region /note="MSTD repeat: matches 1. .341 of consensus"  
27203. .27695  
repeat_region /note="LIP10 repeat: matches 5642. .6165 of consensus"  
27696. .27750  
repeat_region /note="MSTD repeat: matches 341. .394 of consensus"  
27857. .27971  
repeat_region /note="MERSA repeat: matches 84. .189 of consensus"  
27986. .28150  
misc_feature /note="MERSA repeat: matches 3. .183 of consensus"  
28765. .29329  
misc_feature /note="match: GSS: Em:B93338"  
28777. .29151  
repeat_region /note="match: GSS: Em:AQ125799"  
29425. .29482  
repeat_region /note="29 copies 2 mer ac 87% conserved"  
29740. .29805  
repeat_region /note="33 copies 2 mer aa 71% conserved"  
29763. .29807  
repeat_region /note="9 copies 5 mer aaac 82% conserved"  
30205. .30412  
repeat_region /note="MIR repeat: matches 2. .230 of consensus"  
30382. .30732  
misc_feature /note="MERSA repeat: matches 8. .187 of consensus"  
30911. .31378  
misc_feature /note="match: GSS: Em:AQ546895"
```

```
repeat_region 32009. .32080  
gene /note="MIR repeat: matches 132. .206 of consensus"  
32125. .32784  
CDS /gene="ba31k16.1"  
32125. .32784  
/gene="ba31k16.1"  
/note="ba31k16.1 (snoRNA binding domain pseudogene)  
match: proteins: Sw:000567 Tr:Q9YK69 Tr:Q9ZKWO Tr:Q9YD11  
Tr:Q94514 Sw:Q21276 Tr:Q9SGT7 Tr:Q29158 Sw:Q12460  
Tr:Q80401 Tr:Q65334 Tr:Q65335 Tr:Q9U5W4  
/codon_start=1  
/pseudo  
/evidence=not experimental  
repeat_region 34656. .34977  
repeat_region /note="Alus repeat: matches 11. .297 of consensus"  
36947. .36978  
misc_feature /note="16 copies 2 mer aa 87% conserved"  
38608. .39160  
polyA_site /note="match: GSS: Em:AQ353745"  
complement(40339)  
/gene="ELAVL2"  
mRNA complement(join(40339. .43119,43683. .43721,51615. .51840,  
55153. .55306,81257. .81360,112241. .112481,171869. .171992))  
/gene="ELAVL2"  
/note="match: cDNAs: Em:U29088 Em:X85111 Em:M62843  
Em:U17602 Em:U12431 Em:L26405 Em:U17599 Em:U29148  
Em:U17597 Em:D26158 Em:SB3320 Em:D31953 Em:AF176675  
match: ESTs: Em:AA545382 Em:AA081584 Em:A1843112  
Em:AA656788 Em:AA692890 Em:A1145457 Em:AV117363  
Em:AA524152 Em:AA416369 Em:A1028574 Em:AA491988  
Em:AA549705 Em:D61361 Em:AV362071 Em:A1837658 Em:AV094529
```

```
Query Match 3.9% Score 89.2 DB 9 Length 183483;  
Best Local Similarity 64.6% Pred.No. 9.9e-14;  
Matches 133; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
  
QY 891 ctgagccaggctgggcaactctctctgtaaagacagacaggaactattagctc 950  
Db 10360 CTACACACAGAGCTCAGCAAACTTCTCTGTAAAGACAGATGATNAATTTTACGTTT 10419  
  
QY 951 tgtgtgcataatggtctcagtcacaaactactcaatctctgtgtagcagaaagcaat 1010  
Db 10420 TGTGGCCATATWGTCTCTCTGCACAACTACTCCACTCTGCTGTCTGAGGCAAAAGCAGT 10479  
  
QY 1011 tagcaacatactgcacaacacatactgtgaccccatgaaacttatattatgtggtac 1070  
Db 10480 TATAGAAATATCATNAATGAAATGACATGCGTGTGTCCATAAAGTATTATGACAC 10539  
  
QY 1071 ggaacctgaaataatgtcttctt 1096  
Db 10540 TGAATTTGAAATTCATGTAAATTTT 10565
```

Search completed: August 16, 2002, 06:01:15  
Job time: 19139 sec

Mon Aug 19 06:46:31 2002

us-09-515-369b-1.rge

Page 23





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:50:31 ; Search time 97.23 seconds  
(without alignments)  
5775.154 Million cell updates/sec

Title: US-09-515-369B-1

Perfect score: 2286  
Sequence: 1 taatacgcactactataggg.....tgactccacgctgagacg 2286

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/prodata/1/ina/5A.CONB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B.CONB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A.CONB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B.CONB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCYUS.CONB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71.8	3.1	7218	1	US-08-232-463-14
2	67.6	3.0	6235	4	US-09-305-384-5
3	67.6	3.0	6679	4	US-09-305-384-1
4	62.2	2.7	178	4	US-08-991-789A-210
5	62.2	2.7	178	4	US-09-062-451-210
6	52.6	2.3	7218	1	US-08-232-463-14
7	52	2.3	3786	4	US-08-975-762-42
8	52	2.3	3786	4	US-09-295-028-42
9	52	2.3	3786	4	US-09-106-582-42
10	50.6	2.2	502	2	US-08-967-101-91
11	50.6	2.2	502	2	US-08-592-541-91
12	50.6	2.2	502	3	US-09-124-698-91
13	50.6	2.2	502	4	US-09-127-480-91
14	50.6	2.2	502	4	US-08-496-841C-91
15	50.6	2.2	1883	2	US-08-967-101-10
16	50.6	2.2	1883	2	US-08-967-101-10
17	50.6	2.2	1883	2	US-08-592-541-10
18	50.6	2.2	1883	2	US-08-592-541-10
19	50.6	2.2	1883	3	US-08-888-077A-8
20	50.6	2.2	1883	3	US-08-124-698-10
21	50.6	2.2	1883	3	US-09-124-698-10
22	50.6	2.2	1883	4	US-09-127-480-10
23	50.6	2.2	1883	4	US-09-127-480-10
24	50.6	2.2	1883	4	US-08-496-841C-10
25	50.6	2.2	1883	4	US-08-496-841C-10
26	48.2	2.1	80246	4	US-09-078-294-4
27	48.2	2.1	80595	4	US-09-078-294-4

28	45.8	2.0	1734	2	US-08-858-052-2	Sequence 2, App1
29	45.8	2.0	1734	3	US-09-200-284-2	Sequence 2, App1
30	43.6	1.9	460	2	US-08-967-101-48	Sequence 48, App1
31	43.6	1.9	460	2	US-08-592-541-48	Sequence 48, App1
32	43.6	1.9	460	3	US-09-124-698-48	Sequence 48, App1
33	43.6	1.9	460	4	US-09-127-480-48	Sequence 48, App1
34	43.6	1.9	460	4	US-08-496-841C-48	Sequence 48, App1
35	42.8	1.9	59065	4	US-09-813-817-3	Sequence 17, App1
36	42.4	1.9	289	4	US-09-007-005-17	Sequence 17, App1
37	42.4	1.9	289	4	US-09-244-796-17	Sequence 17, App1
38	41.4	1.8	703	4	US-09-313-300-6	Sequence 6, App1
39	41.4	1.8	176373	3	US-09-128-155-17	Sequence 17, App1
40	41.2	1.8	7505	4	US-09-078-294-13	Sequence 13, App1
41	39.8	1.7	350	2	US-08-332-766A-20	Sequence 20, App1
42	39.2	1.7	84495	4	US-09-797-906-3	Sequence 3, App1
43	39.2	1.7	152331	3	US-09-128-155-16	Sequence 16, App1
44	38.6	1.7	43795	3	US-08-742-185-101	Sequence 101, App1
45	38.2	1.7	1732	1	US-08-480-346-2	Sequence 2, App1

#### ALIGNMENTS

RESULT 1  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHREIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-F15  
US-08-232-463-14

Query Match 3.1%; Score 71.8; DB 1; Length 7218;  
Best Local Similarity 4.0%; Pred. No. 9.7e-12;  
Matches 16; Conservative 237; Mismatches 144; Indels 0; Gaps 0;

[illegible]

```

RESULT      2
US-09-305-384-5
; Sequence 5, Application US/09305384
; Patient No. 6242218
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
; FILE REFERENCE: 07236/017001
; CURRENT APPLICATION NUMBER: US/09/305.384
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084.649
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 6235
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-305-384-5

```

Query Match	3.08;	Score 67.6;	DB 4;	Length 6235;
Best Local Similarity	63.7%;	Pred. No. 1.8e-10;		
Matches 135;	Conservative	0;	Mismatches 74;	Indels 3;
				Gaps 2;

Oy 921 aagaagacagaaagaaactttttagacctgtgagcataatgtctcagtaacactac 980  
 11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
 Db 622 aagaagacagaaagaaactttttagacctgtgagcataatgtctcagtaacactac 680  
 11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
 Oy 981 tcatctctcctctgtagcagaagaacttagcaacataatgacaaacataatgta 1040  
 11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
 Db 681 tcgagccgctctctctgacaaagacgacttaaacataacataatgtaattttata 740  
 11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
 Oy 1041 ccccaagaaacttatattatattgtagatcagaaaccgaaataatgtctcttcttga 1100  
 11111 11111 11111 11111 11111 11111 11111 11111 11111 11111  
 Db 741 gacatcgagattgtatttcataatgctt-ttaacatttataaataaacttttataaa 798  
 1101 tttttcccaatcaataaaaaacgtaaaac 1132

Db 799 atttccctaaccatttaaaagtgtaaaagc 830

```

RESULT      3
US-09-305-384-1
: Sequence 1, Application US/09305384
: Patent No. 6242218
: GENERAL INFORMATION:
: APPLICANT: Treco, Douglas A.
: APPLICANT: Heartlein, Michael W.
: APPLICANT: Seiden, Richard F
: TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
: FILE REFERENCE: 07236/017001
: CURRENT APPLICATION NUMBER: US/09/305,384
: CURRENT FILING DATE: 1999-05-05
: EARLIER APPLICATION NUMBER: US 60/084,649
: EARLIER FILING DATE: 1998-05-07
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 6679
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-305-384-1

```

Query Match	3.08;	Score 67.6;	DB 4;	Length 6679;
Best Local Similarity	63.7%;	Pred. No. 1.9e-10;		
Matches 135;	Conservative	0;	Mismatches 74;	Indels 3;
				Gaps 2

[illegible]

RESULT 4  
US-08-991-789A-210  
; Sequence 210, Application US/08991789A  
; Patent No. 6225054

```

; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.

```

Smith, John M.  
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF BREAST CANCER

NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESS: 5004 TD

ADDRESS: Seed LP Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle

CITY: Seattle  
STATE: Washington  
COUNTRY: USA

COMPUTER READABLE FORM:  
ZIP: 98104-7092  
COUNTRY: USA

COMPUTER: IBM PC compatible  
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A

```





```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-967-101-91

Query Match          2.2%: Score 50.6; DB 2; Length 502;
Best Local Similarity 64.3%: Pred. No. 7,6e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgtaagaagaccacagaaactattttagagctctgtgtgcatatggtctcagtcacaa 976
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CTGTAAAAAGCCAGACTG--AATATTTTAACTCTATGCGTCATATGCTCTCCAGGCCAA 321
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 977 ctactcaatctgctctctgtgtgcagaaagcaattagacaacatattgcaacaaacat 1036
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 ACACCACTGCTGTGCTAGCTAGTGTGAAGCAGCAGACAGCAATGTTATTAAACAAGAGG 381
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1037 gtgaaccatgaaaaacttatt 1059
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 GTGGTCACTTCCATGAATGAAGTT 404
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-08-592-541-91
; Sequence 91, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fletcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-91

Query Match          2.2%: Score 50.6; DB 2; Length 502;
Best Local Similarity 64.3%: Pred. No. 7,6e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgttaagaagaccagaaactattttagagctctgtgtgcatatggtctcagtcacaa 976
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CTGTAAAAAGCCAGACTG--AATATTTTAACTCTATGCGTCATATGCTCTCCAGGCCAA 321
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 977 ctactcaatctgctctctgtgtgcagaaagcaattagacaacatattgcaacaaacat 1036
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 ACACCACTGCTGTGCTAGCTAGTGTGAAGCAGCAGACAGCAATGTTATTAAACAAGAGG 381
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1037 gtgaaccatgaaaaacttatt 1059
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 GTGGTCACTTCCATGAATGAAGTT 404
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-08-592-541-91
; Sequence 91, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fletcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-91

```



```

; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
;
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-08-496-841C-91

Query Match          2.2%; Score 50.6; DB 4; Length 502;
Best Local Similarity 64.3%; Pred. No. 7.6e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgtaagaacagacaggaactatttagctctgtgtccatatgttcagtcacaa 976
    ||||||| ||||| | ||||||| ||||| ||||||| |||||
DB 264 CTGTAAAGCCAGACATG--AATATTTTAAAGCTCTATGGGTCAATGCTCCAGGGCAA 321

QY 977 ctactcatctgctctgttagcacgaagaacttagcaacatatgtgcacaacatat 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 322 ACACTCAACTGTGCTACTGTAGTGTGAAGCAGGACAGACATGTATTAAACAGAGAG 381

QY 1037 gtgaacccatgaaaacttattt 1059
    ||| | | | | | | | | | | | | | | | | | | | | |
DB 382 GTGGTCACTTCCATGAAAGTT 404

RESULT 15
US-08-967-101-10/c
; Sequence 10, Application US/08967101
; Patent No. 5840540
;
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Steel Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1883 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-08-967-101-10

Query Match          2.2%; Score 50.6; DB 2; Length 1883;
Best Local Similarity 64.3%; Pred. No. 1.7e-05;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgtaagaacagacaggaactatttagctctgtgtccatatgttcagtcacaa 976
    ||||||| ||||| | ||||||| ||||| ||||||| ||||| |||||
DB 239 CTGTAAAGCCAGACTG--AATATTTTAAAGCTCTATGGGTCAATGCTCCAGGGCAA 182

QY 977 ctactcatctgctctgttagcacgaagaacttagcaacatatgtgcacaacatat 1036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 ACACTCAACTGTGCTACTGTAGTGTGAAGCAGGACAGACATGTATTAAACAGAGAG 122

QY 1037 gtgaacccatgaaaacttattt 1059
    ||| | | | | | | | | | | | | | | | | | | | | |
DB 121 GTGGTCACTTCCATGAAAGTT 99

Search completed: August 16, 2002, 05:14:32
Job time: 15841 sec
```







QY 1076 ccggaataatgtctcttcttgatcttccccaatcaataaagaacaaactac 1135:  
DB 477 TTTTAAATTTCAATATATCTCTTTTGTATATTTTGAACATTTAAATAAAGAACCA 418  
QY 1136 tcttagtcgcaaggttaagccatcttcagcttagcagtgagcagctgagctgtgt 1195  
DB 417 TTTTGTCTCAGCAGGCTATGCCA-----TACAGTCAGCAGCAGATTTGGCTCAT 366  
QY 1196 gacctagcttgcccaatcccgatt 1221  
DB 365 AGGCCACAGTATGCCAACCCCTGTTT 340

RESULT 2  
US-10-027-632-119946/c  
Sequence 119946, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027.632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 119946  
LENGTH: 1007  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-119946

Query Match 3.7%; Score 82.8; DB 7; Length 1007;  
Best Local Similarity 58.0%; Pred. No. 3.1e-12;  
Matches 189; Conservative 0; Mismatches 127; Indels 10; Gaps 2;  
QY 896 cccaggttggcaaaactctctcgttaaaagacagcaggaactctttagctctgtc 955  
DB 655 CCAAGGCTCAGTTAACTTTTCTGTAAGAGCAGAGATTAATTTTAAAGCTTTGAG 596  
QY 956 gacatacgctcagtcacacactcactcctcgtctgtagcaggaagcaatagca 1015  
DB 595 GCCATTAAGCTCTGTCGATCTATCACTTGTGCTGTGTCAGCAATATGACCATAG 536  
QY 1016 acataatgtaacaaacatactgagaccacagaaacttattcttagatcagaaa 1075  
DB 535 ATAAATACATTAATCAATGACATGCTTCAATTAATACT--TTATTTTGAACACTGAA 478  
QY 1076 cctgaataaataatgtctcttcttctgattcttcccaatcatcaaaacgtaaaactac 1135  
DB 477 TTTTAAATTTCAATATATCTCTTTTGTATATTTTGAACATTTAAATAAAGAACCA 418  
QY 1136 tcttagtcgcaaggttaagccatcttcagcttagcagtgagcagctgagctgtgt 1195  
DB 417 TTTTGTCTCAGCAGGCTATGCCA-----TACAGTCAGCAGCAGATTTGGCTCAT 366  
QY 1196 gacctagcttgcccaatcccgatt 1221  
DB 365 AGGCCACAGTATGCCAACCCCTGTTT 340

DB 365 AGGCCACAGTATGCCAACCCCTGTTT 340  
RESULT 3  
US-10-105-299-12107/c  
Sequence 12107, Application US/10105299  
GENERAL INFORMATION:  
APPLICANT: Rosen, et. al  
TITLE OF INVENTION: Human Secreted Proteins  
FILE REFERENCE: PS950  
CURRENT APPLICATION NUMBER: US/10/105,299  
CURRENT FILING DATE: 2002-03-26  
NUMBER OF SEQ ID NOS: 15197  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12107  
LENGTH: 37437  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-105-299-12107

Query Match 3.7%; Score 81.8; DB 7; Length 37437;  
Best Local Similarity 62.9%; Pred. No. 3.1e-11;  
Matches 144; Conservative 0; Mismatches 82; Indels 3; Gaps 1;  
QY 807 actcctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 866  
DB 32330 ACTCCACTTCAATCTTCAATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 32271  
QY 867 ca---tctcactccgcttcccatcctgagccaggttggcaaacctctctgtaa 923  
DB 32270 AATTTGCTCTTGAATGATGCTATGATTAACACAGGGCTGCAAACTTTTCTGAAA 32211  
QY 924 gaaccagcaggaactttaggtctgtgtcctatgctgctgctgctgctgctgctgctg 983  
DB 32210 GAGCCAGATGCTAATATTTTATGCTTGTGGCCGACATTTCTGTGCAACTATTCA 32151  
QY 984 tctcgtcctctgttagcaggaagcaatagcaaatatgtaacaaac 1032  
DB 32150 TTTCTGCTGCTATGATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCT 32102

RESULT 4  
US-10-105-299-12108/c  
Sequence 12108, Application US/10105299  
GENERAL INFORMATION:  
APPLICANT: Rosen, et. al  
TITLE OF INVENTION: Human Secreted Proteins  
FILE REFERENCE: PS950  
CURRENT APPLICATION NUMBER: US/10/105,299  
CURRENT FILING DATE: 2002-03-26  
NUMBER OF SEQ ID NOS: 15197  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12108  
LENGTH: 37442  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-105-299-12108

Query Match 3.6%; Score 81.4; DB 7; Length 37442;  
Best Local Similarity 63.0%; Pred. No. 3.9e-11;  
Matches 143; Conservative 0; Mismatches 81; Indels 3; Gaps 1;  
QY 809 tctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 868  
DB 32329 TCCACTTCAATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCT 32270  
QY 869 ---tctcactccgcttcccatcctgagccaggttggcaaacctctctctgaaaga 925  
DB 32269 ATTGCTCTTGCATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCT 32210

QY	986	tcgcgctctgtagcagcaagcaatctagcaacaatahgcacaaac	1032
Db	32149	tcctcgcctctatgctacaaagcgcctatpaaacatpaaatgacgaac	32103
QY	926	accagcaaggaacattatttagctctcgtgcgaatbgtctacgtcaaacacacac	985
Db	32209	gccagatgcgaattatttagcctttgtagggcgcacacattctcgttgcacatttcatt	32150

RESULT 5  
US-09-629-469A-12215/c  
; Sequence 12215, Application US/09629469A

```

1  APPLICANT: OTA, TOSHIO
2  APPLICANT: ISOGAI, TAKAO
3  APPLICANT: NISHIKAWA, TETSUO
4  APPLICANT: HAYASHI, KOJI
5  APPLICANT: SAITO, KAORO
6  APPLICANT: YAMAMOTO, JUNICHI
7  APPLICANT: ISHII, SHIZUKO
8  APPLICANT: SUGIYAMA, TOMOYASU
9  APPLICANT: WAKAMATSU, AI
10 APPLICANT: NAGAI, KEIICHI
11 APPLICANT: OTSUKI, TETSUJI
12 TITLE OR INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
13 FILE REFERENCE: 084355/0123
14 CURRENT APPLICATION NUMBER: US/09/629,469A
15 CURRENT FILING DATE: 2000-07-28
16 PRIOR APPLICATION NUMBER: JP 1999-248036
17 PRIOR FILING DATE: 1999-07-29
18 PRIOR APPLICATION NUMBER: JP 1999-300253
19 PRIOR FILING DATE: 1999-08-27
20 PRIOR APPLICATION NUMBER: JP 2000-118776
21 PRIOR FILING DATE: 2000-01-11
22 PRIOR APPLICATION NUMBER: JP 2000-183767
23 PRIOR FILING DATE: 2000-05-02
24 PRIOR APPLICATION NUMBER: JP 2000-241899
25 PRIOR FILING DATE: 2000-06-09
26 PRIOR APPLICATION NUMBER: 60/159,590
27 PRIOR FILING DATE: 1999-10-18
28 PRIOR APPLICATION NUMBER: 60/183,322
29 PRIOR FILING DATE: 2000-02-17
30 NUMBER OF SEQ ID NOS: 19025
31 SOFTWARE: PatentIn Ver. 2.1
32 SEQ ID NO 12215
33 LENGTH: 5294
34 TYPE: DNA
35 ORGANISM: Homo sapiens
36 US-09-629-469A-12215

```

Query Match	3.6%	Score 81	DB 5	Length 5294
Best Local Similarity	74.5%	Pred. No. 2.1e-11		
Matches 102	Conservative	0	Mismatches 35	Indels 0
				Gaps 0
QY 894	gaccacggttgggcaaacctcttcctgtlaaagaccagacgaacatattttagctctgt	953		
DB 1915	GACCACGAATTCAGAAACTTTTCTCTAAAGGCCAGAGAGATTTTAGCTTTGT	1856		
QY 954	gtgcacatatgtctcagtcacaaactacatctatctgtcctctgttagcagcaagaacattg	1013		
DB 1855	GGGCCCATGTGGTCTCTCTGCGCAATTACTTACTCTGCGCATATGACAGAAAGCAGCTAC	1796		
QY 1014	caacacatatgtcacaacaa	1030		
DB 1795	AGACATATTGTAGCGCAA	1779		

RESULT 6  
US-09-918-995-12893/C  
; Sequence 12893, Application US/09918995  
; GENERAL INFORMATION:

```

1  APPLICANT Hyseq, Inc.
2  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
3  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
4  FILE REFERENCE: 20411-756
5  CURRENT APPLICATION NUMBER: US/09/918,.995
6  CURRENT FILING DATE: 2001-07-30
7  PRIOR APPLICATION NUMBER: US/09/235,076
8  PRIOR FILING DATE: 1999-01-20
9  NUMBER OF SEQ ID NOS: 38054
10 SOFTWARE: FastSeq for Windows Version 3.0
11 SEQ ID NO 12893
12 LENGTH: 503
13
14 TYPE: DNA
15
16 ORGANISM: Homo sapiens
17
18 FEATURE:
19 NAME/KEY: misc-feature
20 LOCATION: (1)..(503)
21 OTHER INFORMATION: n = A,T,C or G
22
23 US-09-918-.995-12893

```

Query Match	3.58;	Score 79.4;	DB 5;	Length 503;
Best Local Similarity	58.4%;	Pred. No. 2e-11;		
Matches 157;	Conservative	0;	Mismatches 111;	Indels 1;
			Gaps	1;

OY	887	atctcggagcccaagtttggagaaactcttcctgttaaagaacacagagacattttatg	946
Db	459	atgcccttccaattccatgsgtgcaggaactctgtatgctgscacagtgatgtaactgattg	400
OY	947	gactctgtgcatatgtgtctcaagtaaacatcaactcattctgtcctgttagaagaaga	1006
Db	399	gcctcttgtaggcacacacattctgtgtttagacnctnctnactgtctgccctctgacatgacg	340
OY	1007	caattgacacaaatattgtcaacaacatatgtgagcccaagaacattctttatatatg	1066
Db	339	cagtcacacagatattatgtgaattgacatgacatgagatgctggccctctttttaaacaattttcatt	280
OY	1067	atctcggaacactgt-aaataatgtcttctcttctgtatcttttccccaactcatataaaca	1125
Db	279	acattatgacacatgctcacaataattattattcttttggaatttttttttcagtcattttaaagcta	220
OY	1126	taaaactactctcttagtgcgaagtttaa	1154
Db	219	taaaaaacacaaacatgcaaaaaaaactttaa	191

```

RESULT 7
US-09-785-276A-55453/c
: Sequence 55453, Application US/09785276A
: GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007B
CURRENT APPLICATION NUMBER: US/09/785,276A
CURRENT FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0

```

```

? SEQ ID NO 55453
? LENGTH: 627
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 561..582
? OTHER INFORMATION: n = A,T,C or G
? US-09-785-276A-55453

```

Query Match	3.58;	Score 79.4;	DB 5;	Length 627;
Best Local Similarity	67.7%;	Pred. No. 2.2e-11;		
Matches 128; Conservative	0;	Mismatches 56;	Indels 5;	Gaps 1.

Qy	907	caaaaccttcctgtaaaagaaacagaaacggagacatttttagctctcgtgctgcatatggtc	966
Db	476	CGAAGCTTTCCCTGTAAGGCGCAAGGATTAATTTTCAGACTTGTTAAGCCAAACATC	417
Qy	967	tcaagtaacaactacatcatctctgcctctctgtagaagcaagcaattagcaacaatagca	1020
Db	416	TCTGTCAAAATATATCAACTGTGGCATTTGTTGGCAAGGACGACGACAAACATATACAA	357
Qy	1027	acaaacaatctgtgac-----cccaatgaacactttatctatctatgtaacaggaacctgaa	1081
Db	356	ACAAAGTGAACCTGGCTGTGTTCCAAATAAACTTTATTTACAACAACTAGAGGTAAGCCTGAT	297
Qy	1082	aataatgc 1090	
Db	296	GATCGAGTTC 288	

RESULT 8  
US-10-027-632-151080  
Sequence 151080, Application US/10027632  
General Information

```

? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
?
? FILE REFERENCE: 108627.123
?
? CURRENT APPLICATION NUMBER: US/10/027.632
?
? CURRENT FILING DATE: 2002-04-30
?
? PRIOR APPLICATION NUMBER: US 60/218,006
?
? PRIOR FILING DATE: 2000-07-12
?
? PRIOR APPLICATION NUMBER: US 60/198,676
?
? PRIOR FILING DATE: 2000-04-20
?
? PRIOR APPLICATION NUMBER: US 60/193,483
?
? PRIOR FILING DATE: 2000-03-29
?
? PRIOR APPLICATION NUMBER: US 60/185,218
?
? PRIOR FILING DATE: 2000-02-24
?
? PRIOR APPLICATION NUMBER: US 60/167,363
?
? PRIOR FILING DATE: 1999-11-23
?
? PRIOR APPLICATION NUMBER: US 60/156,358
?
? PRIOR FILING DATE: 1999-09-28
?
? PRIOR APPLICATION NUMBER: US 60/146,002
?
? PRIOR FILING DATE: 1999-08-09
?
? NUMBER OF SEQ ID NOS: 325720
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO 151080
?
? LENGTH: 749
?
? TYPE: DNA
?
? ORGANISM: Human
?
? US-10-027-632-151080

```

Query Match	3.58;	Score 79;	DB 7;	Length 749;
Best Local Similarity	72.08;	Pred. No. 3.1e-11;		
Matches 103; Conservative	0;	Mismatches 40;	Indels 0;	Gaps 0

```
Oy      885 ccacgacctgcaccaggtctggcacaaccccttccttgtaagaacccagcacgaatactattt 944
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     372 cctggttttagccagcggtggcgcaaacctcatgctgtataaaagccagatgctgaataatttt 431
```

Oy	945	agcgtctggtgtcaataagtccttaagcaaacatcatctctgcctctgttagcaaga	1004
Db	432	agcgtctgtggtgcatttgagcctctgtcaaacactaactaagctctgcgtctgttagcaaa	491
Oy	1005	agcaattagcaacaatattgttaa	1027
Db	492	accagccacagacaataatgtctaa	514

RESULT 9  
US-10-027-632-151081

```

Sequence: 51081, Application: US/10027632
GENERAL INFORMATION:
  APPLICANT: Wang, David G.
  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
  TITLE OF INVENTION: Polymorphisms In the Human Genome
  FILE REFERENCE: 108827.129
  CURRENT FILING DATE: 2002-04-30
  PRIOR APPLICATION NUMBER: US 60/218,006
  PRIOR FILING DATE: 2000-07-12
  PRIOR APPLICATION NUMBER: US 60/198,676
  PRIOR FILING DATE: 2000-04-20
  PRIOR APPLICATION NUMBER: US 60/193,483
  PRIOR FILING DATE: 2000-03-29
  PRIOR APPLICATION NUMBER: US 60/185,218
  PRIOR FILING DATE: 2000-02-24
  PRIOR APPLICATION NUMBER: US 60/167,363
  PRIOR FILING DATE: 1999-11-23
  PRIOR APPLICATION NUMBER: US 60/156,358
  PRIOR FILING DATE: 1999-09-28
  PRIOR APPLICATION NUMBER: US 60/146,002
  PRIOR FILING DATE: 1999-08-09
  NUMBER OF SEQ ID NOS: 325720
  SOFTWARE: FastSeq For Windows Version 4.0
  SEQ ID NO 151081
  LENGTH: 749
  TYPE: DNA
  ORGANISM: Human
US-10-027-632-151081

```

Query Match	3.58;	Score 79;	DB 7;	Length 749;
Best Local Similarity	72.08;	Pred. No. 3.1e-11;		
Matches 103; Conservative	0;	Mismatches 40;	Indels 0;	Gaps 0;

OY	885	ccatgctgcgacccacgagctctgggcaaacctctctctctgaagaagacacgaacactcttt	944
Dd	372	cctggttttaggcacagagtgagggcgaactctcatgtgtaaagaacgaagtgtgtaaatcttt	431
OY	945	aggctctgtgtggtccatctatgctctcagcacaactcactctatctctgctctgtatgacacga	1004
Dd	432	aggactctgtgggcacatcttggtgcctctgtcaacaactactacgtctgtgcctgtatgagcga	491
OY	1005	agcgaattgacacaaatattgtctaa	1027
Dd	492	accagaccacagacacattatgttaaa	514

RESULT 10  
US-10-027-632-7748/C  
; Sequence 7748, Application US/10027632

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7748
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-7748

```

```

Query Match 3.5%; Score 78.2; DB 7; Length 754;
Best Local Similarity 64.1%; Pred. No. 5.1e-11;
Matches 116; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

```

```

QY 892 tggaccaggttggcaaacctctctgttaagaacagaggaactatttagctct 951
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
DB 239 TGGAGCAGGCGCTTGGCAAACTTTCTGTAAAGGCAAGATGTAATATTATAGGCTTT 180
QY 952 gtgtccatattgtctcaagcaactactcatctctgtctgttagcagaagaacatt 1011
    ||| | : | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 179 GTGACCTTTCAGGCTCTGTGACAGACTACTCAACTCTGCCATTGTGTAAAGAAACAGCC 120
QY 1012 agcaacaatattgtcaacaacataatgtaccacataaacttattattatgatgacg 1071
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
DB 119 ACAGACAAATACGAAAAAAGAAACATGTTGTATTCACAACTTTATTATTAATAACA 60
QY 1072 g 1072
DB 59 G 59

```

```

RESULT 11
US-10-104-047-364
; Sequence 364, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NOVEL full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 364
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-104-047-364

```

```

Query Match 3.4%; Score 77.2; DB 7; Length 2960;
Best Local Similarity 74.6%; Pred. No. 1.8e-10;
Matches 97; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
DB 2470 ggaagcagggcagcaaacatttctgttaagggccagataagataatttgggcttgg 2529
QY 953 ttgtccatattgtctcaagcaactactcatctctgtcctctgttagcagaagaacatta 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2530 tgggacctatgtctctgtcaacaagcttcaactctgtctgttaaaaaaagcagta 2589
QY 1013 gcaacaatat 1022

```

```

DB 2590 gacatatgt 2599

```

```

RESULT 12
US-10-027-632-186700/c
; Sequence 186700, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186700
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-186700

```

```

Query Match 3.4%; Score 76.8; DB 7; Length 575;
Best Local Similarity 63.3%; Pred. No. 1.1e-10;
Matches 133; Conservative 1; Mismatches 73; Indels 3; Gaps 1;

```

```

QY 891 ctggaccaggttggcaaacctctctgttaagaacagaggaactatttagctc 950
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 384 CTAGACGAGAGTAGGCAAACTTTCTGTAAAGGCGAGATGTAATATTATAGGATT 325
QY 951 ttgtccatattgtctcaagcaactactcatctctgtcctctgttagcagaagaacatt 1010
    || | : | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 324 TGGAGCTTTCAGGCTCTGTGACAGACTACTCAACTCTGCCATTGTGTAAAGAAATCAGC 265
QY 1011 tagcaacaatattgtcaacaacataatgtaccacataaacttattattatgatgac 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 264 CACAGATATATATATAAACAATGAGCATGCTGTGT---TCCAGTAACACTTATGTTCAG 208
QY 1071 ggaacccgaaataatattgtctctcttga 1100
DB 207 TGAATCTGATTTTAACAGCTTTCATGTCA 178

```

```

RESULT 13
US-10-027-632-142654/c
; Sequence 142654, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

```

```

; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142654
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142654
```

```

Query Match      3.4%; Score 76.6; DB 7; Length 538;
Best Local Similarity 64.7%; Pred. No. 1.2e-10;
Matches 130; Conservative 0; Mismatches 69; Indels 2; Gaps 1;
```

```

QY 901 gtgggcaaacctctcctgtaagaacacagacagaaacatttaggcctgtgtgcaat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 214 GGTGGCAAACTTTCTCTTAAGGACGAGTAAACATTTTAGGTTTGCAGGCAC 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atggtcagtcacacactacatctctgctctgtagcagaaagcaattagcaacat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 154 ATGGCTCTGTACACTACTCTACCTCTGCACTGTAGCCTCAACAGCACAGACAT 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacatgtgacccacatgaaacttattt-attatgatacggaaacct 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 94 ATTTAATTAGCGAAGCGTGGCGCTCTCCATTAAACTTCATTGTGTACTGAATTT 35
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1079 gaaataatgtctctcttctt 1099
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 34 GAATTTAATTAATTTTCATG 14
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 14
US-09-785-276A-55307/c
; Sequence 55307, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55307
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-785-276A-55307
```

```

Query Match      3.4%; Score 76.6; DB 5; Length 608;
Best Local Similarity 74.0%; Pred. No. 1.3e-10;
Matches 97; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```

QY 901 gtgggcaaacctctcctgtaagaacacagacagaaacatttaggcctgtgtgcaat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 GTTGGCAAACTTTCTCTTAAGGACGAGTAAACATTTTAGGTTTGCAGGCAT 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atggtcagtcacacactacatctctgctctgtagcagaaagcaattagcaacat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 AAGGCTCTGTCAACTATTCATCTCTGCACTTTTGTGAGAAAGCACATGACAT 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacaa 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 GTATAACAAA 53
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 15
US-10-027-632-233742/c
; Sequence 233742, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233742
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-233742
```

```

Query Match      3.4%; Score 75.6; DB 7; Length 556;
Best Local Similarity 65.9%; Pred. No. 2.3e-10;
Matches 108; Conservative 1; Mismatches 55; Indels 0; Gaps 0;
```

```

QY 901 gtgggcaaacctctcctgtaagaacacagacagaaacatttaggcctgtgtgcaat 960
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 472 GTTGGCAAACTTTCTCTTAAGGACGAGTAAACATTTTAGGTTTGCAGGCAT 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 atggtcagtcacacactacatctctgctctgtagcagaaagcaattagcaacat 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 412 ATGGCTCTATCACAACACTGTGCTCAGCCACTGTAGCAAAAAGCATTTGTAACAT 353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 atgtcaacaacatagtacccacatgaaacttatttattat 1064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 352 ATGTAAGTGAATGAACATGGGTGTCTCAATAAATTTAATTT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Search completed: August 16, 2002, 05:57:46
Job time: 11760 sec
```



1

2

3

4



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 07:54:01 ; Search time 5721.79 Seconds  
(without alignments)  
8469.757 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240

Perfect score: 2240  
1 taatcgcctactactagg.....aaatcccaagaagcttgc 2240

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/US006.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US006.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US007.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
10: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
11: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
12: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
13: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
14: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
15: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
16: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
17: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
18: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
19: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
20: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
21: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
22: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
23: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
24: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
25: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
26: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
27: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
28: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
29: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
30: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
31: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
32: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
33: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
34: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
35: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
36: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
37: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
38: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
39: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
40: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
41: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
42: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*  
43: /cgn2\_6/ptodata/1/pna/US008.COMB.seq:\*

Result No.	Score	Query length	ID	Description
1	2240	100.0	US-09-515-369B-1	Sequence 1, Appl
2	420	18.8	US-09-528-409-108247	Sequence 108247, Appl
3	420	18.8	US-09-933-524-108247	Sequence 108247, Appl
4	420	18.8	US-09-933-524-108247	Sequence 108247, Appl
5	93	4.2	PCT-US01-15674A-421	Sequence 421, Appl
6	93	4.2	US-09-573-080A-421	Sequence 421, Appl
7	87	3.9	US-60-245-228-200	Sequence 200, Appl
8	87	3.9	US-60-258-227-80	Sequence 80, Appl
9	86	3.8	PCT-US01-15674A-420	Sequence 420, Appl
10	86	3.8	US-09-573-080A-420	Sequence 420, Appl
11	86	3.8	US-60-243-468-358	Sequence 358, Appl
12	83.2	3.7	US-60-226-176-1401	Sequence 1401, Appl
13	83.2	3.7	US-60-233-468-1401	Sequence 1401, Appl
14	83.2	3.7	US-60-313-371-1401	Sequence 1401, Appl
15	83	3.7	US-60-212-664-256	Sequence 256, Appl
16	82.8	3.7	US-09-634-306B-119945	Sequence 119945, Appl
17	82.8	3.7	US-09-634-306B-119946	Sequence 119946, Appl
18	82.4	3.7	US-60-212-358-69	Sequence 69, Appl
19	81.8	3.7	US-60-044-030-1633	Sequence 1633, Appl
20	81.8	3.7	US-60-172-360-1441	Sequence 1441, Appl
21	81.8	3.7	US-60-185-216-346	Sequence 346, Appl
22	81.8	3.7	PCT-US01-01354-38977	Sequence 38977, Appl
23	81.8	3.7	US-09-764-903-38977	Sequence 38977, Appl
24	81.8	3.7	US-09-764-903-38977	Sequence 38977, Appl
25	81.8	3.7	US-10-092-399-38977	Sequence 38977, Appl
26	81.8	3.7	PCT-US01-01354-38978	Sequence 38978, Appl
27	81.4	3.6	US-09-764-903-38978	Sequence 38978, Appl
28	81.4	3.6	US-09-950-083-9957	Sequence 9957, Appl
29	81.4	3.6	US-10-092-399-38978	Sequence 38978, Appl
30	81.4	3.6	US-60-226-176-1603	Sequence 1603, Appl
31	81.2	3.6	US-60-226-176-1603	Sequence 1603, Appl

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

```

C 32 81.2 3.6 161395 62 US-60-233-468-1603 Sequence 1603, Ap
C 33 81.2 3.6 161395 70 US-60-313-371-1603 Sequence 1603, Ap
C 34 80.8 3.6 32768 59 US-60-208-965-68 Sequence 68, Appl
C 35 79.4 3.5 503 16 US-09-235-076-12893 Sequence 12893, A
C 36 79.4 3.5 503 16 US-09-248-797-34353 Sequence 34353, A
C 37 79.4 3.5 503 17 US-09-332-782-12893 Sequence 12893, A
C 38 79.4 3.5 503 29 US-09-737-223-12893 Sequence 12893, A
C 39 79.4 3.5 503 34 US-09-925-564-34353 Sequence 34353, A
C 40 79 79 749 24 US-09-634-306B-151080 Sequence 151080,
C 41 79 79 749 24 US-09-634-306B-151081 Sequence 151081,
C 42 78.4 3.5 312283 1 PCT-US00-27620-91 Sequence 91, Appl
C 43 78.2 3.5 754 24 US-09-634-306B-7748 Sequence 7748, Ap
C 44 78.2 3.5 118794 63 US-60-243-468-310 Sequence 310, Appl
C 45 78.2 3.5 118794 63 US-60-243-742-53 Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-515-369B-1
; Sequence 1, Application US/09515369B
; GENERAL INFORMATION:
; APPLICANT: Paul B. Fisher
; APPLICANT: Malavi T. Medireddi
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; FILE REFERENCE: 34611 070050.1685
; CURRENT APPLICATION NUMBER: US/09/515, 369B
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Human
US-09-515-369B-1

Query Match 100.0%; Score 2240; DB 19; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 taatcgactacataaggcgctgacatcgatccatttgaaacccaggtctgctgct 60
DB 1 taatcgactacataaggcgctgacatcgatccatttgaaacccaggtctgctgct 60
QY 61 ccaaaagctgtactacataactagatctcaactgtatgtttggccaaagtttcttagttct 120
DB 61 ccaaaagctgtactacataactagatctcaactgtatgtttggccaaagtttcttagttct 120
QY 121 ctctctgacatctctctctgaaagtaataatgctatgaatcgcacgtgaggtctgagcc 180
DB 121 ctctctgacatctctctctgaaagtaataatgctatgaatcgcacgtgaggtctgagcc 180
QY 181 aggcacatgtttgctctgacataatcattatgattctctctcagaagaatgagct 240
DB 181 aggcacatgtttgctctgacataatcattatgattctctctcagaagaatgagct 240
QY 241 actacagatcccaagtgtaacctctgagcgaagcgaagtgtatcctacatgacct 300
DB 241 actacagatcccaagtgtaacctctgagcgaagcgaagtgtatcctacatgacct 300
QY 301 gtccagcctgctcttaacagctatccacatgacctgacctccgccatctgcaga 360
DB 301 gtccagcctgctcttaacagctatccacatgacctgacctccgccatctgcaga 360
QY 361 cagtagtctaggaattctgctgctgctgctgctgctgctgctgctgctgctgct 420
DB 361 cagtagtctaggaattctgctgctgctgctgctgctgctgctgctgctgctgct 420
QY 421 gctgtctctctgctccacatcactatattccagcaatccacatgctgcttcttctg 480
DB 421 gctgtctctctgctccacatcactatattccagcaatccacatgctgcttcttctg 480
```

```

DB 421 gctgtctctctgctccacatcactatattccagcaatccacatgctgcttcttctg 480
QY 481 tctcatcactgctctcttgacatcttattatcctatagtagtgtaggttcttctgta 540
DB 481 tctcatcactgctctcttgacatcttattatcctatagtagtgtaggttcttctgta 540
QY 541 gccccaataaccacatggtggtggaaggtggtggtggtggtggtggtggtggtggt 600
DB 541 gccccaataaccacatggtggtggaaggtggtggtggtggtggtggtggtggtggt 600
QY 601 gctactctctgaggtgagtagacatcggtcgctcccaagaacaaagattaggtctg 660
DB 601 gctactctctgaggtgagtagacatcggtcgctcccaagaacaaagattaggtctg 660
QY 661 gctatagccaagcagactgctggtccagaggtatgcaaaagatatttctgcttaaga 720
DB 661 gctatagccaagcagactgctggtccagaggtatgcaaaagatatttctgcttaaga 720
QY 721 aataaacaacactgtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 780
DB 721 aataaacaacactgtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 780
QY 781 gactctgccaaggtggtgttctactcaatctctctcttcttcttcttcttcttcttct 840
DB 781 gactctgccaaggtggtgttctactcaatctctctcttcttcttcttcttcttcttct 840
QY 841 aggcagttatctgtgcccccaagctacattcttactcccggttccatgctgtagccag 900
DB 841 aggcagttatctgtgcccccaagctacattcttactcccggttccatgctgtagccag 900
QY 901 gttggccaactctctccgttaagaagcagaagaactattttaggtctggtgctgct 960
DB 901 gttggccaactctctccgttaagaagcagaagaactattttaggtctggtgctgct 960
QY 961 atgtgtctagtcacaactactatctctctgctctgtatgacgaagaattagcaacaat 1020
DB 961 atgtgtctagtcacaactactatctctctgctctgtatgacgaagaattagcaacaat 1020
QY 1021 atgtcaacaacacatgtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1080
DB 1021 atgtcaacaacacatgtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1080
QY 1081 aataatgtcttcttcttctgatttcttcccaatcatataaagaagtataaactactt 1140
DB 1081 aataatgtcttcttcttctgatttcttcccaatcatataaagaagtataaactactt 1140
QY 1141 ggtcgcaaggttaagcacttctcagcttagcaggtgacagctggaattctgctgtgac 1200
DB 1141 ggtcgcaaggttaagcacttctcagcttagcaggtgacagctggaattctgctgtgac 1200
QY 1201 acaattgccaatccctgtagtcccaaaaatgattctcctcagggatgtaggcaaatctat 1260
DB 1201 acaattgccaatccctgtagtcccaaaaatgattctcctcagggatgtaggcaaatctat 1260
QY 1261 ggaaggtgctggtatataaagaaggttaagaagcaltcagaacattccagagcgtagcac 1320
DB 1261 ggaaggtgctggtatataaagaaggttaagaagcaltcagaacattccagagcgtagcac 1320
QY 1321 atggcagggcctcttaactgtaacctatgtagtattcatgtttcatgtagagattctgcaag 1380
DB 1321 atggcagggcctcttaactgtaacctatgtagtattcatgtttcatgtagagattctgcaag 1380
QY 1381 acaagaattctcctcaaaactagtagtctgtagagcgtgcttggagaacaacatgctctg 1440
DB 1381 acaagaattctcctcaaaactagtagtctgtagagcgtgcttggagaacaacatgctctg 1440
QY 1441 atgacctcaactggtgacatgtaggaatctagagctgtagtctgtagctgtagagtag 1500
DB 1441 atgacctcaactggtgacatgtaggaatctagagctgtagtctgtagctgtagagtag 1500
QY 1501 gtcaagaactcttgaacgtgctgtagcaggtctgtagaacaacacagcgtctgtgtccctga 1560
DB 1501 gtcaagaactcttgaacgtgctgtagcaggtctgtagaacaacacagcgtctgtgtccctga 1560
```

QY	1561	ggtcagagagccacacagggccaaagactatagcccccacagacacaaagccctcccaagctttctg	1620
QY	1561	ggtctccagagccacacagggccaaagactatagcccccacagacacaaagccctcccaagctttctg	1620
QY	1621	gactctgcctcgtcgtatcatcttccattctcctcgtgtttccagaacctaaagaaagagcac	1680
Db	1621	gactctgcctcgtcgtatcatcttccattctcctcgtgtttccagaacctaaagaaagagcac	1680
QY	1681	attttggttgatgatatataaacccctagggagacatvggttagctgcacatgcagaaacatc	1740
Db	1681	attttggttgatgatatataaacccctagggagacatvggttagctgcacatgcagaaacatc	1740
QY	1741	ctcaactctccctgcccccttgatgataaagaagagagtgacttcaagaattattcttcgcgt	1800
Db	1741	ctcaactctccctgcccccttgatgataaagaagagagtgacttcaagaattattcttcgcgt	1800
QY	1801	ggacactctgccagacatgaatagggacatcatattgaattattttgcagaaagaggac	1860
Db	1801	ggacactctgccagacatgaatagggacatcatattgaattattttgcagaaagaggac	1860
QY	1861	tgtggtatgtgtgcacagatataatgtatttacaatggtacagagtttccaaagacccct	1920
Db	1861	tgtggtatgtgtgcacagatataatgtatttacaatggtacagagtttccaaagacccct	1920
QY	1921	gtgtgtgttttgcctttgtttatatacacttgggaacaaattttaaattatacatgcag	1980
Db	1921	gtgtgtgttttgcctttgtttatatacacttgggaacaaattttaaattatacatgcag	1980
QY	1981	agactgtagagcagagaagaagcttaagagaacttgcacctgcgcacacagccagtggttagagcc	2040
Db	1981	agactgtagagcagagaagaagcttaagagaacttgcacctgcgcacacagccagtggttagagcc	2040
QY	2041	tgaactcaaaacccaaagtctcatctcaactcaactcaaggagctcttcccatcgtctatgtc	2100
Db	2041	tgaactcaaaacccaaagtctcatctcaactcaactcaaggagctcttcccatcgtctatgtc	2100
QY	2101	cttaagatgattggtgtgcacagcaatggaagtattctctagaaagaagatgaccaaatttc	2160
Db	2101	cttaagatgattggtgtgcacagcaatggaagtattctctagaaagaagatgaccaaatttc	2160
QY	2161	ctttctccacctccctcttllttctccacacctcccccacagcccccataatatgtccc	2220
Db	2161	ctttctccacctccctcttllttctccacacctcccccacagcccccataatatgtccc	2220
QY	2221	aaatctcacaagaagccttgc 2240	
Db	2221	aaatctcacaagaagccttgc 2240	

```

US-09-528-409-108247/c
Sequence 108247, Application US/09528409
GENERAL INFORMATION:
APPLICANT: Dramanac, Radoje T.
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Dickson, Mark
APPLICANT: Jones, Lee W.
TITLE OF INVENTION: Novel Nucleic Acid Seq
TITLE OF INVENTION: From Various Libraries
FILE REFERENCE: 774
CURRENT APPLICATION NUMBER: US/09/528,409
CURRENT FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,453
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 116231
SOFTWARE: Hy-patent.pl Version 3.1
SEQ. ID NO. 108247
LENGTH: 426
TYPE: DNA
ORGANISM: Homo sapiens
US-09-528-409-108247

```

Query Match	18.8%;	Score 420;	DB 19;	Length 426;
Best Local Similarity	100.0%;	Pred. No. 1.6e+108;		
Matches 420;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	556	gqvggaaagggggagatvggggaaabaaagtcgcctgtgvgggcgtgtgcctactcttggaa	615
Db	425	GGTGGGAAGCGCGAGTGGGGGAAGAGTCCGTGTGGGGCTGTGCTACTCTTCTGGAG	366
QY	616	gtaagactcgggcctccagaacaagaattcaagctltgtgtgcactatagccaag	675
Db	365	GTAAGACTCGGGCCCTCCAGGAAGAAAGACTTCAGCTGTGGCAGGTATAGCCAGCAG	306
QY	676	actcgtggccaggatctgcaaaagagatattgttctgctaagaanaaacaactga	735
Db	305	ACGCTGGCCAGGATTCGCAAGAGATTTTCTTTCTTAAGAAATATACAACTGA	246
QY	736	gtatagagatlgagggaggggtgtgtgtgtgcagagaaattgggaagatctgtccaaaggt	795
Db	245	GATAGAGTGGAGGAGGGGTGTGTGTGGCCAGAGAAATTTGGCAAGGTGTGCCAAGGT	186
QY	796	gtgtctactcaactcctcttcttcttctatctccactgactgtgagagcaattatcctgt	855
Db	185	GTGTTCTACTACTCTCTCTCTTTCTTTCTTTCACCTCCACTGAGCTSGAAGCATTTATCTGT	126
QY	856	cccccacgtcaacttctactctccgtttcccaigtctggaccacagtttgggcaactct	915
Db	125	CCCCACGCTCAATTCCTACTCCCGTTTCCATCCTCGGAGCCGAGCTTGGCAAACTTT	66
QY	916	ctgtgaagaacccagacaggaactattttagctctgttgcacatatygtctcaagtcaca	975
Db	65	CTGTGAAGAACCAGACAGCACTATTTTTAAAGCTCTGTGTGCCATATAGTCTCAAGTACA	6

```

? RESULT 3
? US-09-933-524-108247/c
? Sequence 108247, Application US/09933524
? GENERAL INFORMATION:
? APPLICANT: Dymenac, Radoje T.
? APPLICANT: Labat, Ivan
? APPLICANT: Stache-Cralin, Birgit
? APPLICANT: Dickson, Mark
? APPLICANT: Jones, Lee W.
? TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
? FILE REFERENCE: 774
? CURRENT APPLICATION NUMBER: US/09/933,524
? CURRENT FILING DATE: 2001-08-20
? PRIOR APPLICATION NUMBER: 09/528,409
? PRIOR FILING DATE: 2000-03-17
? NUMBER OF SEQ ID NOS: 116231
? SOFTWARE: HY-patent.pl Version 3.1
? SEQ ID NO 108247
? LENGTH: 426
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-933-524-108247

```

	Query Match	18.8%	Score 420	DB 35	Length 426
	Best Local Similarity	100.0%	Pred. No.	1.6e-108	
	Matches 420	Conservative	0	Mismatches 0	Indels 0
				Gaps 0	
Qy	556	ggttggaagggggaagttggtggaagaagatgcctgttgaggcttcgccacttcggaag	615		
Db	425	ggttggaagggggaagttggtggaagaagatgcctgttgaggcttcgccacttcggaag	366		
Qy	616	gtaagaactgcggccctccaggaacaagaatcagagctgtgtgcagctctaaagcag	675		
Db	365	gtaagaactgcggccctccaggaacaagaatcagagctgtgtgcagctctaaagcag	306		
Qy	676	actgcgtgcgcgaaggaattgcgaagaaggaatttgccttaagaaataaacaacatcga	735		



Query Match 3.88; Score 86; DB 1; Length 224;



[illegible]

```

RESULT 13
US-60-233-468-1401
: Sequence 1401, Application US/60233468
: GENERAL INFORMATION:
: APPLICANT: Ring, Huijun Z.
: APPLICANT: Malsen, Gareth
: APPLICANT: Townley, David
: APPLICANT: Morris, MacDonald
: APPLICANT: Valdes, Ana
: TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes
: FILE REFERENCE: GX-0013-2 P
: CURRENT APPLICATION NUMBER: US/60/233,468
: CURRENT FILING DATE: 2000-09-18
: NUMBER OF SEQ ID NOS: 2488
: SOFTWARE: PERL Program
: SEQ ID NO 1401
: LENGTH: 179040
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: GB:AC022732_3
US-60-233-468-1401

```

Query Match	3.7%	Score	83.2	DB	62	Length	179040
Best Local Similarity	63.2%	Pred	No. 1e-10				
Matches	144	Conservative	0	Mismatches	83	Indels	1
							Gaps
QY	879	cgttcccatgctgagcccaaggttggcgaaactctccctgtaaaagaccagacgaac	938				
Db	5023	cttggagcaagcttaagagacaggygttcaacaacctttttgtaaaagaccagaagtaa	5082				
QY	939	tattttagagctctgtgcatatgctctcagtcacaactactactctctgcctctgag	998				
Db	5083	tatttcaagcttttgcacagcccatatgctctctgcgcgaactgttcaacctctccagggatag	5142				
QY	999	cacgaaagcaattagacaaatafatgcaacaaacatatgtaacccacatgaaacattatt	1058				
Db	5143	caccaagagcagtcataagacaata-citcaagcaanagagctgctactataataatttatt	5201				
QY	1059	tattatgtagtagaagaaacctgaaataaagctctctcttttgatttttt	1106				
Db	5202	tatgactggaatttgtaatttcataactttatagttctctctcttttttt	5249				

```

RESULT 14
US-60-313-371-1401
; Sequence 1401, Application US/60313371
; GENERAL INFORMATION:
; APPLICANT: Ring, Huijun Z.
; APPLICANT: Malsen, Gareth
; APPLICANT: Townley, David
; APPLICANT: Morris, Macdonald
; TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated with ADME Genes
; FILE REFERENCE: GX-0013-5 P
; CURRENT APPLICATION NUMBER: US/60/313.371

```

```
? CURRENT FILING DATE: 2001-08-16
? NUMBER OF SEQ ID NOS: 2447
? SOFTWARE: PERL Program
? SEQ ID NO 1401
? LENGTH: 179040
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: GB:AC022732_3
JS-60-313-371-1401
```

Query Match	3.78;	Score 83.2;	DB 70;	Length 179040;
Best Local Similarity	63.78;	Need No	10-10.	

[illegible]

```

RESULT 15
US-60-212-664-256/c
; Sequence 256, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Spier, Gene
; APPLICANT: Greenberg, Simon
; APPLICANT: Rabkin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: C1000687
; CURRENT APPLICATION NUMBER: US/60/212,664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 256
; LENGTH: 163174
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(163174)
; OTHER INFORMATION: n = A,T,C or G
; US-60-212-664-256

```

	Query Match	Best Local Similarity	3-7% 65.2%	Score 83	DB 60	Length 163174
	Matches 122	Conservative	0	Mismatches 65	Indels 0	Gaps 0
QY	903	tgggcaaacctcttcgcgtgaagaaacacagagacattttgctcgtctgtgccaat	962			
DB	92341	TCGGCAACCTTTCCTGTGTAAGGTGACATAGTAATAATTTTATGGCTTGTGGCCACAG	92282			
QY	963	ggtctcagtcacacactcactcctctctgctctttagaagaagaacatttagcacacatat	1022			
DB	92281	AGGCTGTGTGCACAAATTCGCCAATCTTACTGTGTGTGAGCCAAAGCAGGCTAATTAATATGT	92222			

```
OY 1023 gtcacacacacatctgtgaccacatgaaacatttatctatcgtatcaggaacctgaa 1082
    |||||
Db 92221 GTAAATGAATTAACGTGSCAATGTTCGCAAAAACCTTTATTATGAGAGCTGAATTCAT 92162
OY 1083 ataatgt 1089
    |||||
Db 92161 ATAAATT 92155
```

Search completed: August 16, 2002, 08:16:27  
 Job time: 16584 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 03:39:47 ; Search time 3499.63 seconds  
(without alignments)  
8638.965 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatcagcaccatagagg.....aatctccacaagccttc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pin:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	3.6	480	10	BI262402
2	80.6	3.6	354	10	T58770
3	80.4	3.6	390	10	I44398
4	78.4	3.5	521	12	A0695956
5	78.2	3.5	435	9	AW183569
6	78.2	3.5	459	9	AW125442
7	78.2	3.5	628	12	A0378792
8	77.4	3.5	393	9	A1128823
9	77.4	3.5	411	10	W76494
10	77.2	3.5	680	12	AG154552
11	77.2	3.4	400	9	A1926274
12	77	3.4	322	9	A1240516
13	77	3.4	386	9	AA443938
14	77	3.4	393	9	AA444117
15	77	3.4	417	9	A1208768
16	76.8	3.4	509	10	BI061153
17	76.6	3.4	327	12	B98848

18	76.2	3.4	499	12	A0277351
c 19	76.2	3.4	644	12	A0039307
c 20	76.2	3.4	5529	12	AF101969
c 21	76	3.4	573	12	A0633901
c 22	75.6	3.4	530	12	A0508112
c 23	75.6	3.4	628	12	A0020407
c 24	75.6	3.4	799	10	BG535645
c 25	75.2	3.4	488	9	AV604853
c 26	74.6	3.3	348	12	A0035111
c 27	73.6	3.3	640	12	AG093736
c 28	73.2	3.3	381	9	AM574982
c 29	73.2	3.3	512	12	A0768179
c 30	73	3.3	519	12	A0029411
c 31	72.8	3.2	302	9	A1559884
c 32	72.8	3.2	427	12	A0815705
c 33	72.2	3.2	312	10	BG230513
c 34	72	3.2	416	9	AM835797
c 35	72	3.2	764	10	BI915216
c 36	71.8	3.2	2032	11	AF161549
c 37	71.4	3.2	794	10	BF983586
c 38	71.2	3.2	327	9	A1344682
c 39	71.2	3.2	363	9	A1358986
c 40	71.2	3.2	372	9	A1370636
c 41	71.2	3.2	380	9	A1244892
c 42	71.2	3.2	380	9	A1298789
c 43	71.2	3.2	388	9	A1291781
c 44	71.2	3.2	393	9	A1214952
c 45	71.2	3.2	394	9	A1439755

## ALIGNMENTS

RESULT 1  
BI262402 480 bp mRNA linear EST 17-JUL-2001  
LOCUS 60293383671 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5086033 3,  
DEFINITION mRNA sequence.  
ACCESSION BI262402  
VERSION BI262402.1 GI:14822582  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
1 (bases 1 to 480)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1993)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cga@bbs-rt@mail.nih.gov](mailto:cga@bbs-rt@mail.nih.gov)  
Tissue procurement: Dr. Louis Staudt, M.D., Ph.D. cDNA Library  
Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILMU at:  
<http://image.llnl.gov>  
Plate: LMC1846 row: g column: 10  
High quality sequence start: 27  
High quality sequence stop: 480.  
Location/Qualifiers

## FEATURES

source  
1..480  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5086033"  
/clone\_lib="NIH-MGC\_99"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
BcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GCCACGAC(G). Size-selected >500bp for average insert size



```
Db 149 CCAGAGACAATACATAGCAAGGAGACTATGTGTCACATAGACTTTATTATATA 208
QY 1070 cggaaacctgaaataatgctc 1091
Db 209 AACAAATGTCACACTATAGTTT 230

RESULT 4
ACCESSION A0695956 521 bp DNA linear GSS 06-JUL-1999
DEFINITION HS_2160_A2_P03_T7C CIT Approved Human Genomic Sperm Library D Homo
ACCESSION A0695956
VERSION A0695956.1 GI:5386204
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 521)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2160 row: K column: 6
Seq primer: T7
Class: BAC ends
High quality sequence stop: 521.
Location/Qualifiers
1..521
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Plate=2160 Col=6 Row=K"
/sex="male"
/note="Organ: Sperm; Vector: pBelobAC11; BAC Clones in
E-coli DH10B"

BASE COUNT 150 a 128 c 92 g 148 t 3 others
ORIGIN
Query Match 3.5%; Score 78.4; DB 12; Length 521;
Best Local Similarity 73.5%; Pred. No. 4.2e-09;
Matches 100; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 906 gcaaaccttcctgtaagaacacagacagacattttaggcctctgycatagtc 965
Db 290 GCAAACTTTTCTGTAAAGGCGCAGATAGTAATATGTGATGCTTGCAGACATATGTT 349
QY 966 ctgcagcaaacactacatcctctgctctgtagcaagaacaattagcaacaatagtc 1025
Db 350 CTCGTTCACACTACTAGCTCTCGGCTTGACCAAAAGCAGCCATAGCAATATGCA 409
QY 1026 aacaaactatgtgac 1041
Db 410 GACAAATGATGTGAC 425

RESULT 5
ACCESSION AM183569 435 bp mRNA linear EST 18-NOV-1999
LOCUS
```

```
DEFINITION xj76h04.x1 soares_nfl_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2663191.3', mRNA sequence.
ACCESSION AM183569
VERSION AM183569.1 GI:6452083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 435)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 420.
Location/Qualifiers
1..435
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2663191"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site.1: Not I; Site.2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL15W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
1.W.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 149 a 82 c 60 g 144 t
ORIGIN
Query Match 3.5%; Score 78.2; DB 9; Length 435;
Best Local Similarity 59.6%; Pred. No. 4.5e-09;
Matches 201; Conservative 0; Mismatches 108; Indels 28; Gaps 3;

QY 901 gtgggcaaacctctctgtaagaacacagacagacattttaggcctctgycat 960
Db 108 GGTGAGCAAACTTTTCTGTAAAGGCGCAGATAGTAATATGTGATGCTTGCAGACAT 167
QY 961 atggtctcagctcaaacactacatcctctgctctgtagcaagaacaattagcaacaat 1020
Db 168 TTGCTCTTTCTGTAACACTACTGCTGTATACATATAGCAAAAGCTGCCAGACAT 227
QY 1021 atgtcaacaacaatcagtcgaccccatgaacctatctat--tatggatcagaacc 1077
Db 228 ATGTAAACAATAGAGGCTGACTGTAAATATGTATTTATGACACTGAAATTTGAACT 287
QY 1078 tgaataaatgctctct-----ttgatcttllccccaatcaatcaataaa 1121
Db 288 TCAAAATCATTTTCATGCTTCCCAAAACATTCTTCTTGACATTTTTCACCATTTTAA 347
QY 1122 aacgtaaaacactccttagagtcgcaaggttaagccatctcagcttagcagtgagcagc 1181
Db 348 AAGCTAAAAACATTTCTTAGAGTGACAGGCTGTGAAAAACT-----AGTCATAGGC 398
QY 1182 tggatttgcttgacctacagttggccaatccctg 1218
Db 399 TTAATCTGGCCACAGCCATAGTTCGCCAATCCTG 435

RESULT 6
```

AI125442 459 bp mRNA linear EST 28-OCT-1998  
LOCUS qd89b02.x1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:1736619  
DEFINITION 3', mRNA sequence.  
ACCESSION AI125442  
VERSION AI125442.1 GI:3593956  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 459)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLND at:  
[www.bio.livl.gov/bdtp/image/image.html](http://www.bio.livl.gov/bdtp/image/image.html)  
Insert Length: 592 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham.  
Location/Qualifiers  
1..459  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="1736619"  
/clone\_lib="Soares\_testis\_NHT"  
/sex="male"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer [5'  
TGTTCACCAATCTGAGAGGAGCGCCGCAATTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Col2, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
BASE COUNT 159 a 88 c 61 g 151 t  
ORIGIN  
Query Match 3.5%; Score 78.2; DB 9; Length 459;  
Best Local Similarity 59.1%; Pred. No. 4.6e-09;  
Matches 204; Conservative 0; Mismatches 113; Indels 28; Gaps 3;  
QY 901 gttgggcaaacctctccgttaagaagacaggaactatttaggtctgtgtgcac 960  
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||||| ||||| |||||  
DB 105 GGTGAGCAAACTTTTCTGTAAAGGCCATACATATTTTCAGTCTGTGACACCAT 164  
QY 961 atgtctcagtcacactactatctctgtcctctgtacagcaagaacatagcaacat 1020  
||||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 165 TTGGCTTTGCTGTAACACTCACTGTCTGCTACATTAACAAAGCTGCCACACACAT 224  
QY 1021 atgtcaacaacatagtcagcccatgaaaacttactat---tatgatacggaaac 1077  
||||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 225 ATGTAAACAAATGAGGTGACTGTAAATAATGTTATTATGACACCTGAATTTGAAC 284  
QY 1078 tgaataatgtcttctc-----ttgatcttccccaatcaatataa 1121  
||||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 285 TCAATTCATTTCATGCTTCCCAAAACATTCTACTTTTGACATATTTTCAACCATTTAA 344  
QY 1122 aacgtaaaaactacttagtcgcaaggttaagccattctcagcttagcagtgcaagc 1181  
||||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 345 AAAGTAAACCATCTTAGTGTGACAGGCTGTAGAAAACCT-----AGTCATAGGC 395

QY 1182 tggattgctgtgactacagttgagcaatccctgattcccaa 1226  
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 396 TTAACTGTGGCCACAAAGCCATATGTCGCAATGCTGTACACATA 440  
RESULT 7  
QY 396 TTTAATCTGGCCACAAAGCCATATGTCGCAATGCTGTACACATA 440  
LOCUS AQ378792/C  
DEFINITION RPCI11-151B24.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-151B24,  
DNA sequence.  
ACCESSION AQ378792  
VERSION AQ378792.1 GI:4349815  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 628)  
AUTHORS Zhao,S., Adams,M.D., Nierman,M., Malek,J., de Jong,P. and Venter  
,J.C.  
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
JOURNAL Unpublished (1997)  
COMMENT Other-GSSs: RPCI11-151B24.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [jheetlgr.org](mailto:jheetlgr.org)  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieterdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genetics (info@resgen.com). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1..628  
/organism="Homo sapiens"  
/db\_xref="GDB:757647"  
/db\_xref="taxon:9606"  
/clone\_image="RPCI-11-151B24"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBAC3.0; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"  
BASE COUNT 173 a 133 c 136 g 185 t 1 others  
ORIGIN  
Query Match 3.5%; Score 78.2; DB 12; Length 628;  
Best Local Similarity 72.1%; Pred. No. 5.1e-09;  
Matches 101; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 892 tggagccaggttgggcaaacctctccgttaagaagacaggaactatttaggtct 951  
||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| ||||| |||||  
DB 572 TAGGCCAGGGGTTGGCACTTTTCTGTAAAGGCCAGATAGGAATATTTCAATTNT 513  
QY 952 gtgtgcacatagtcacactactatctctgtcctctgtacagcaagaacat 1011  
||||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 512 GTGGGCCCTTATGCTCTCTGTACAAACATTAATTCAGCTGTATGACATGAAGCAGCC 453  
QY 1012 agcaacatagtcacaaa 1031  
||| ||||| ||||| ||||| ||| ||||| ||| ||||| ||||| |||||  
DB 452 ACAGACAGTAGTAACAAA 433  
RESULT 8  
AT128823





ACCESSION AI240516  
 VERSION AI240516.1 GI:3835913  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 322)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.  
 Insert Length: 735 Std Error: 0.00  
 Seq primer: -40UP from 61bco  
 High quality sequence stop: 315.  
 Location/Qualifiers  
 1..322  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1844499"  
 /clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
 /lab\_host="DH10B"  
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung Nbh19W, testis NHT, and B-cell) NCI-CGAP (GCB) were mixed and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 728096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 90 a 72 c 67 g 93 t  
 ORIGIN  
 Query Match 3.4%; Score 77; DB 9; Length 322;  
 Best Local Similarity 71.6%; Pred. No. 8.4e-09;  
 Matches 101; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 QY 891 ctggaccaggttgggcaaaccttccgtgaagaacagacaggaactatttagctc 950  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 178 CTAGACACAGAGGTTGGCAACTCTACTGTATTGTCCAGATAGTAATTTTAGGTT 119  
 QY 951 tctgttcacataatgtctcagtcacaaactatctctcctctgtagcagaagaacat 1010  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 118 TGTGGGGCATAGGGCTCTCTATCTCACTACTCACTCTGCTGTGACATGAAGTAGC 59  
 QY 1011 tagcaacaataatgtcaacaacaa 1031  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 58 CATAGATGTTAAGTCAACACAA 38  
 RESULT 13  
 AA443938  
 LOCUS 386 bp mRNA linear EST 03-JUN-1997  
 DEFINITION z551c03.s1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:757156  
 3', mRNA sequence.  
 ACCESSION AA443938  
 VERSION AA443938.1 GI:2156613  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 386)  
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, V., Moore, B., Schellenberg, K., Steploc, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
 WashU-Merck EST Project 1997  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [estewatson.wustl.edu](mailto:estewatson.wustl.edu)  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.  
 Seq primer: -41m3 fwd. ET from Amersham  
 High quality sequence stop: 321.  
 Location/Qualifiers  
 1..386  
 /organism="Homo sapiens"  
 /db\_xref="GDB:597767"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:757156"  
 /clone\_lib="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer (5' TGTTACATCTGTAAGTGGAGCGGCGCCCAATTTTCTTTTCTTTT 3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 111 a 72 c 68 g 135 t  
 ORIGIN  
 Query Match 3.4%; Score 77; DB 9; Length 386;  
 Best Local Similarity 63.0%; Pred. No. 8.8e-09;  
 Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
 QY 881 ttcccatcctcgtgaccaggttgggcaaaccttccgtgaagaacagacagaacta 940  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 139 TTTTTCCTATCTAGAACCAAGGTTGGCAACTTTTCTATTAAGGCGCTGATTAATAA 198  
 QY 941 tttagacctctgttcacataatgtcacaacaacataatgtgaccccatgaaactat 1000  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 199 TTTACAGCTTTGAGGGCTTCACATGCTCTGTGACACTTCAATTCGCCATTATAGTA 258  
 QY 1001 cgaagcaattagcaacaataatgtcacaacaacataatgtgaccccatgaaactat 1060  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 259 CAGAACTGCCACAGACATAGCTCAACAATAAGATGAGTGTGTGCCAATTAACACTT 318  
 QY 1061 tcatggata 1069  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 319 ATTTCCAA 327  
 RESULT 14  
 AA444117/c  
 LOCUS 393 bp mRNA linear EST 03-JUN-1997  
 DEFINITION z551c03.r1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:757156  
 5', mRNA sequence.  
 ACCESSION AA444117  
 VERSION AA444117.1 GI:2156792  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 393)  
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 05:33:21 ; Search time 485.09 seconds

(without alignments)  
7928.195 Million cell updates/sec

Title: us-09-515-369b-1\_COPY\_1\_2240

Perfect score: 2240

Sequence: 1 taatcagctacactataggg.....aaatccacaagaagccttgc 2240

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2240	100.0	2286	22	AAH26595
2	81.8	3.7	37437	22	AAK84165
3	81.4	3.6	37442	22	AAK84166
4	81	3.6	5294	22	AAH14600
5	77.6	3.5	26591	22	AAH13613
6	75.8	3.4	481	22	AAK74761
7	73	3.3	1837	22	AAH04430
8	71.6	3.2	628	22	AAH98619
9	71.4	3.2	286	21	AAA41353

C	10	71.2	3.2	407	21	AAK27788	Human secreted pro
C	11	70	3.1	123219	23	AAH88703	Human DNA sequence
C	12	69.8	3.1	50000	20	AAK23517	Human Kidney amino
C	13	69.4	3.1	8029	22	AAH9747	Human protein enco
C	14	69.4	3.1	46765	24	AAH99306	DNA encoding Aldeh
C	15	69.4	3.1	47319	22	AAK64813	Human immune/haema
C	16	69.4	3.1	47319	22	AAK72230	Human immune/haema
C	17	69.4	3.1	236303	22	AAH11614	Human genomic DNA
C	18	68.6	3.1	325791	22	AAH433104	Human Oestrogen re
C	19	68.2	3.0	97662	22	AAH83908	Genomic sequence o
C	20	68	3.0	31140	22	AAH08065	Human ovarian and
C	21	68	3.0	31140	22	AAH06791	Human reproductive
C	22	67.8	3.0	17700	22	AAK68945	Human immune/haema
C	23	67.6	3.0	2515	22	AAH17444	Human granulocyte
C	24	67.6	3.0	6235	21	AAK29169	Human G-CSF genom
C	25	67.6	3.0	6679	21	AAK29170	Human granulocyte
C	26	67.6	3.0	6679	22	AAH17443	Human granulocyte
C	27	66.6	2.9	289	22	AAH39009	Novel human diagno
C	28	66.4	3.0	466	22	AAH33971	Human colon cancer
C	29	66.4	3.0	655	22	AAH17197	Human breast cancer
C	30	66	2.9	871	22	AAH04280	Human CDNA clone (
C	31	66	2.9	1712	22	AAH13903	Human CDNA sequenc
C	32	66	2.9	2291	22	AAH17522	Human CDNA sequenc
C	33	66	2.9	7496	22	AAK70490	Human immune/haema
C	34	66	2.9	7496	22	AAK74614	Human immune/haema
C	35	66	2.9	10225	22	AAH99021	Human excretory re
C	36	66	2.9	10225	22	AAH63371	Human kidney relat
C	37	66	2.9	24292	22	AAH15939	Human nervous syst
C	38	66	2.9	28731	22	AAK64887	Human immune/haema
C	39	66	2.9	28731	22	AAK69666	Human immune/haema
C	40	66	2.9	28731	22	AAK70491	Human immune/haema
C	41	66	2.9	28731	22	AAK74615	Human immune/haema
C	42	66	2.9	28731	22	AAK84942	Human immune/haema
C	43	65.6	2.9	872	22	AAH98871	Human immune/haema
C	44	65.4	2.9	519	22	AAH35469	Human EST-derived
C	45	64.8	2.9	15044	22	AAH36290	Human musculoskele
							Human cardiovascular

## ALIGNMENTS

RESULT	1
ID	AAH26595 standard; DNA; 2286 BP.
XX	XX
AC	AAH26595;
XX	XX
DT	12-NOV-2001 (first entry)
XX	XX
DE	Human melanoma differentiation associated gene-7 (Mda-7) promoter.
KW	Melanoma differentiation associated gene-7; Mda-7; promoter; human;
KW	neuroblastoma; astrocytoma; glioblastoma multiforme;
KW	cervical cancer; breast cancer; colon cancer; prostate cancer;
KW	osteosarcoma; chondrosarcoma; tumour; therapy; ds.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	1..2240
FT	promoter
FT	/*tag= a
FT	138..146
FT	/*tag= b
FT	/*bound_moiety= "C/EBP"
FT	446..454
FT	/*tag= c
FT	/*bound_moiety= "C/EBP"
FT	773..777
FT	/*tag= d
FT	/*bound_moiety= "C/EBP"
FT	801..807
FT	/*tag= e
FT	/*bound_moiety= "AP-1"

```

FT      protein_bind      1109..1113
FT      /*tag= f
FT      /bound_molecety= "C/EBP"
FT      protein_bind      1207..1214
FT      /*tag= g
FT      /bound_molecety= "AP1"
FT      protein_bind      1209..1214
FT      /*tag= h
FT      /bound_molecety= "C/EBP"
FT      protein_bind      1530..1537
FT      /*tag= i
FT      /bound_molecety= "AP1"
FT      protein_bind      2115..2121
FT      /*tag= j
FT      /bound_molecety= "AP1"
FT      protein_bind      2126..2133
FT      /*tag= k
FT      /bound_molecety= "C/EBP"
FT      TATA_signal      2210..2214
FT      /*tag= l
FT
FT      WO200164921-A1.
FT
FT      07-SEP-2001.
FT
FT      28-FEB-2001; 2001WO-US06782.
FT
FT      29-FEB-2000; 2000US-0515369.
FT
FT      (UYCO ) UNIV COLUMBIA NEW YORK.
FT
FT      Fisher PB, Madiredi MT:
FT      WPI: 2001-565508/63.
FT
FT      Melanoma differentiation associated gene-7 promoter capable of
FT      treating cancer comprises directing transcription of heterologous
FT      coding sequence encoding tumor suppressor polypeptide positioned
FT      downstream, useful for treating cancer.
FT
FT      Claim 1; Fig 6A-B; 132pp; English.
FT
XX      CC      The present sequence is that of the promoter region of the human
XX      CC      melanoma differentiation associated gene-7 (Mda-7). It includes
XX      CC      2240 bp from the 5' flanking region of the Mda-7 gene. The
XX      CC      promoter was isolated from a human placental genomic library
XX      CC      using a PCR-based method. The Mda-7 promoter exhibits melanocyte
XX      CC      tissue specificity and can only be activated in the targeted
XX      CC      tissue, i.e. the skin. Therefore, a gene of interest driven by
XX      CC      the Mda-7 promoter will be differentially expressed in these cells,
XX      CC      minimizing systemic toxicity. A recombinant expression construct
XX      CC      in which the human Mda-7 promoter is operably linked to a coding
XX      CC      sequence encoding a tumor suppressor protein is claimed. The
XX      CC      tumor suppressor is preferably p21, retinoblastoma protein or p53.
XX      CC      A host cell comprising the expression construct is also claimed,
XX      CC      and is preferably a tumor cell selected from a melanoma,
XX      CC      neuroblastoma, astrocytoma, glioblastoma multiforme, cervical
XX      CC      cancer, breast cancer, lung cancer or prostate cancer cell.
XX      CC      The promoter preferably comprises nucleotides 1-2240 of the present
XX      CC      sequence. A pharmaceutical composition including the recombinant
XX      CC      expression construct is used in a claimed method of treating
XX      CC      melanoma, neuroblastoma, astrocytoma, glioblastoma multiforme,
XX      CC      cervical cancer, breast cancer, colon cancer, prostate cancer,
XX      CC      osteosarcoma, chondrosarcoma or a cancer of the central nervous
XX      CC      system.
XX
SO      Sequence 2286 BP; 547 A; 592 C; 510 G; 637 T; 0 other:

```

```

Query Match      100.0%; Score 2240; DB 22; Length 2286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1      taatagactactataggcgctgactgatcaaccttttgaacccaggtctgacct 60
QY      |||
DB      1      taatacgaactcaacataaggcgctgactgatcaaccttttgaacccaggtctgacct 60
QY      61      ccaaaagcttctactacatactagatcttcaactgtatgtggccaaggtctcctagtctc 120
QY      |||
DB      61      ccaaaagcttctactacatactagatcttcaactgtatgtggccaaggtctcctagtctc 120
QY      121      ctcttgaacctctctcttgaagtaatagtctatgaataactcagagagctgaagccc 180
QY      |||
DB      121      ctcttgaacctctctcttgaagtaatagtctatgaataactcagagagctgaagccc 180
QY      181      aggcacatgttctgctgaactatccatgttatatgatctctccccaagagctgaagct 240
QY      |||
DB      181      aggcacatgttctgctgaactatccatgttatatgatctctccccaagagctgaagct 240
QY      241      actacagatcccaaggtgtacacctgaagccagccaaggtgtatccatgaactcagctct 300
QY      |||
DB      241      actacagatcccaaggtgtacacctgaagccagccaaggtgtatccatgaactcagctct 300
QY      301      gtccagagctgacctttaacagatccatccacactgacctcccgccatctgcaga 360
QY      |||
DB      301      gtccagagctgacctttaacagatccatccacactgacctcccgccatctgcaga 360
QY      361      cagtagcttaggatttcaagctgcccgggggtcatttccctcagctctcctgtctta 420
QY      |||
DB      361      cagtagcttaggatttcaagctgcccgggggtcatttccctcagctctcctgtctta 420
QY      421      gctgtctctgctcccccacactatcaactcagacactcctcagctctcctctctctg 480
QY      |||
DB      421      gctgtctctgctcccccacactatcaactcagacactcctcagctctcctctctctg 480
QY      481      tctcatcactgctctctgtacatcttcatctcatagtagttagttaggggtctctgtaac 540
QY      |||
DB      481      tctcatcactgctctctgtacatcttcatctcatagtagttagttaggggtctctgtaac 540
QY      541      gccctaaatccacatggtgggaagggaggttgggggaagaggtctgtgggctgt 600
QY      |||
DB      541      gccctaaatccacatggtgggaagggaggttgggggaagaggtctgtgggctgt 600
QY      601      gctactctctgagaggtagaactctgggacctcaggaacaaaggtatcaggtgtgtgca 660
QY      |||
DB      601      gctactctctgagaggtagaactctgggacctcaggaacaaaggtatcaggtgtgtgca 660
QY      661      gctatagccaagcagactctgtgccaaggtatgtgcaaggaatttctgtgcttaagaa 720
QY      |||
DB      661      gctatagccaagcagactctgtgccaaggtatgtgcaaggaatttctgtgcttaagaa 720
QY      721      aataaacacacctgagtagatgagatggaggggtgtgtgtgccaagagattgggaa 780
QY      |||
DB      721      aataaacacacctgagtagatgagatggaggggtgtgtgtgccaagagattgggaa 780
QY      781      gagctgccaaggtgtgtgttctactcaactctctctcttcttcaatcctcagctgtg 840
QY      |||
DB      781      gagctgccaaggtgtgtgttctactcaactctctctcttcttcaatcctcagctgtg 840
QY      841      aggcagttatctctgtccccaagatcaatctctactccggttcccaagcttgaacag 900
QY      |||
DB      841      aggcagttatctctgtccccaagatcaatctctactccggttcccaagcttgaacag 900
QY      901      gtctggcaaacctctcctgttaagaacagacaggaactatctttagctctgtgccc 960
QY      |||
DB      901      gtctggcaaacctctcctgttaagaacagacaggaactatctttagctctgtgccc 960
QY      961      atgtctcagttacacactactcatctctgtcctctgttagcagaaagcatttagacaac 1020
QY      |||
DB      961      atgtctcagttacacactactcatctctgtcctctgttagcagaaagcatttagacaac 1020
QY      1021      atgttaacaaataatgtgaccccatgaaactatattatattatgatacaggaacctgga 1080
QY      |||
DB      1021      atgttaacaaataatgtgaccccatgaaactatattatattatgatacaggaacctgga 1080
QY      1081      aataatgtcttcttcttgaatttctcccaatcatataaaaaagtaaaactactctta 1140

```

|||||  
Db 1081 aataaagctcttcttggatttttccccaatcatlaaaaaagtaaaactacttta 1140  
Qy 1141 ggtcgagaggttaagccattctcagcttagagtgagagcgctggaatttgcttgagcc 1200  
Db 1141 ggtcgagaggttaagccattctcagcttagagtgagagcgctggaatttgcttgagcc 1200  
Qy 1201 acagcttgccaatccctcgatccccaatgtatctcccaaggagtggtggcaataactat 1260  
Db 1201 acagcttgccaatccctcgatccccaatgtatctcccaaggagtggtggcaataactat 1260  
Qy 1261 ggggaagtgctggatlaaacaagaglttaagaagcatcagaacattccagaagcgctagcac 1320  
Db 1261 ggggaagtgctggatlaaacaagaglttaagaagcatcagaacattccagaagcgctagcac 1320  
Qy 1321 atgccaaggctcttaactaactacatctgatactgttttaaggaagcttggaag 1380  
Db 1321 atgccaaggctcttaactaactacatctgatactgttttaaggaagcttggaag 1380  
Qy 1381 acaagaattccctcaaaactagaagtcgagagctgcttctgggaaacaactgctctgctg 1440  
Db 1381 acaagaattccctcaaaactagaagtcgagagctgcttctgggaaacaactgctctgctg 1440  
Qy 1441 atgcccctcaactgggacatgtagaattcagagctgagctgcttctgtagctggaatagg 1500  
Db 1441 atgcccctcaactgggacatgtagaattcagagctgagctgcttctgtagctggaatagg 1500  
Qy 1501 gtcagaagctcttgaactgacctgagcaactgtaacacacagctgctgctccctgagct 1560  
Db 1501 gtcagaagctcttgaactgacctgagcaactgtaacacacagctgctgctccctgagct 1560  
Qy 1561 ggttcagaagccacaagaagcgcaagaactagcccacaagaagcagcgctcccaactcttcg 1620  
Db 1561 ggttcagaagccacaagaagcgcaagaactagcccacaagaagcagcgctcccaactcttcg 1620  
Qy 1621 ggtctgctcactgctgaattccttcttcttcttcttcttcttcttcttcttcttcttctg 1680  
Db 1621 ggtctgctcactgctgaattccttcttcttcttcttcttcttcttcttcttcttcttctg 1680  
Qy 1681 atttggctgagtgatataataacctagaagcacaaggtagctgcaatgctcagaacacctc 1740  
Db 1681 atttggctgagtgatataataacctagaagcacaaggtagctgcaatgctcagaacacctc 1740  
Qy 1741 ctcaactcctgcccctgagatgaattaaagagaggtacttaacaggtattcttctgctg 1800  
Db 1741 ctcaactcctgcccctgagatgaattaaagagaggtacttaacaggtattcttctgctg 1800  
Qy 1801 ggaactactgcccagaactagatagggacatcatlgaattatttggacagaagaagagac 1860  
Db 1801 ggaactactgcccagaactagatagggacatcatlgaattatttggacagaagaagagac 1860  
Qy 1861 tggctgtaactgctcagaactaataatgtaattacatgtagacagaagttaccacaacacctct 1920  
Db 1861 tggctgtaactgctcagaactaataatgtaattacatgtagacagaagttaccacaacacctct 1920  
Qy 1921 ggtgtgttttgccttgattatataactgggacaatttttaaaattatatacagcag 1980  
Db 1921 ggtgtgttttgccttgattatataactgggacaatttttaaaattatatacagcag 1980  
Qy 1981 agactgcagcgagagaagctaaagaagctgcccctgcccacacagcagctgtgtagagcc 2040  
Db 1981 agactgcagcgagagaagctaaagaagctgcccctgcccacacagcagctgtgtagagcc 2040  
Qy 2041 tgaactccaacccaagctcactatccactcagagagctgcttcccatcgctgattgctc 2100  
Db 2041 tgaactccaacccaagctcactatccactcagagagctgcttcccatcgctgattgctc 2100  
Qy 2101 cttaaagtgatgggtgaactaggaatgaatgttctcaggaagaacatgacaaatttcc 2160  
Db 2101 cttaaagtgatgggtgaactaggaatgaatgttctcaggaagaacatgacaaatttcc 2160  
Qy 2161 cttctccacctccctcttcttctccacccctccccaaccaccccaataatagccc 2220  
Db 2161 cttctccacctccctcttcttctccacccctccccaaccaccccaataatagccc 2220

Db 2161 cttctccacctccctcttcttctccacccctccccaaccaccccaataatagccc 2220  
Qy 2221 aaatctccacaagcctgac 2240  
Db 2221 aaatctccacaagcctgac 2240  
RESULT 2  
AAK84165/c  
ID AAK84165 standard; DNA; 37437 BP.  
XX  
AC AAK84165;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38977.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX  
PN MO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225279.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227109.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.



RESULT 3  
AAK84166/c  
ID AAK84166 standard; DNA: 37442 BP.  
XX  
AC AAK84166;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human Immune/haematopoietic antigen genomic sequence SPO ID NO:38978.  
XX  
KM Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX  
XX MO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189674.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000: 2000US-0251030.  
PR 05-DEC-2000: 2000US-0251988.  
PR 05-DEC-2000: 2000US-0256719.  
PR 06-DEC-2000: 2000US-0251479.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251859.  
PR 08-DEC-2000: 2000US-0251989.  
PR 11-DEC-2000: 2000US-0254097.  
PR 05-JAN-2001: 2001US-0253678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI: 2001-483426/52.  
XX  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX  
PS Disclosure: SEQ ID NO 38978: 3071pp + Sequence Listing: English.  
XX  
XX  
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
XX  
SQ Sequence 37442 BP: 10361 A: 6715 C: 6860 G: 13506 T: 0 other:  
  
Query Match 3.6%; Score 81.4; DB 22; Length 37442;  
Best Local Similarity 63.0%; Pred. No. 2.1e-12;  
Matches 143; Conservative 0; Mismatches 81; Indels 3; Gaps 1;  
  
QY 809 tctctctcttcttccacgtgagctgagcagttatccctgctcccaagtcacac 868  
DB 33329 TCCACTTCATCTTCTTGTGCTGCTCTTCCAGTATATATTCCTAAACAAATAAAA 32270  
QY 869 -----tctctctccgcttccacgtgagcagcttgggaacacttctctgaaga 925  
DB 32269 ATTGTCTCTTTCGACATGCTAGATATCTTAACACAGGGGCTCAGCAAACTTTTCTGAAAAA 32210  
QY 926 accagacaggaactatttagctctgtgtgcacatagttctcagtcacaactactacac 985  
DB 32209 GCCAGATGGTAAATATTAGGCTTTGTGGCCAGACATTCCTGTTGCAACTATTCATT 32150  
QY 986 tctgctctgtgacaggaagcaatagcaacaatatgtcaacaac 1032  
DB 32149 TCTGCTGCTATAGTACAAAGACAGCTATAAACAATAGTAGCAAAAC 32103  
  
RESULT 4  
AAH14600/c  
ID AAH14600 standard; cDNA: 5294 BP.  
XX  
AC AAH14600;  
XX  
DT 26-JUN-2001 (first entry)

XX  
XX  
DE Human cDNA sequence SEQ ID NO:12215.  
XX  
XX  
KM Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
XX  
OS Homo sapiens.  
XX  
PN Ep1074617-A2.  
XX  
XX  
PD 07-FEB-2001.  
XX  
XX  
PF 28-JUL-2000: 2000EP-0116126.  
XX  
XX  
PR 29-JUL-1999: 99JP-0248036.  
XX  
PR 27-AUG-1999: 99JP-0300253.  
XX  
PR 11-JAN-2000: 2000JP-0118776.  
XX  
PR 02-MAY-2000: 2000JP-0183767.  
XX  
PR 09-JUN-2000: 2000JP-0241899.  
XX  
XX  
PA (HELI-) HELIX RES INST.  
XX  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakematsu A, Nagai K, Otsuki T;  
XX  
XX  
DR WPI: 2001-318749/34.  
XX  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX  
PS Claim 8: SEQ ID 12215: 2537pp + CD ROM; English.  
XX  
XX  
CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
XX  
SQ Sequence 5294 BP: 1637 A: 878 C: 1099 G: 1680 T: 0 other:  
  
Query Match 3.6%; Score 81; DB 22; Length 5294;  
Best Local Similarity 74.5%; Pred. No. 9.1e-13;  
Matches 102; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
QY 894 gaccaggttggaacactcttcgttaagaacagacagaaactatttagctctgt 953  
DB 1915 GACCAGGATCAGAAAACTTTTCTGTAAAGGCCAAGAGATTAAGTTTGGT 1856  
QY 954 gtgcacatagtctcagtcacaactactcattctgctctgtgacaggaacattag 1013  
DB 1855 GGGCCATATGGCTCTGTGGCAATTAATTAACTGCGCATATAGCAGAAAGCCTAC 1796  
QY 1014 caacacatgtcaacaa 1030



DB 1795 AGACATATGTACGCA 1779

RESULT 5  
AAL36313/c  
ID AAL36313 standard; DNA; 26591 BP.  
XX  
AC AAL36313;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.  
XX  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW cytotoxic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US01338.  
PF  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  

---

PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.



PR	14-SEP-2000	2000US-02323.99
PR	14-SEP-2000	2000US-02324.00
PR	14-SEP-2000	2000US-02324.01
PR	14-SEP-2000	2000US-02330.63
PR	14-SEP-2000	2000US-02330.65
PR	14-SEP-2000	2000US-02330.66
PR	21-SEP-2000	2000US-02344.23
PR	21-SEP-2000	2000US-02344.27
PR	21-SEP-2000	2000US-02349.97
PR	25-SEP-2000	2000US-02354.98
PR	25-SEP-2000	2000US-02354.98
PR	27-SEP-2000	2000US-02358.34
PR	27-SEP-2000	2000US-02358.36
PR	29-SEP-2000	2000US-02363.27
PR	29-SEP-2000	2000US-02363.67
PR	29-SEP-2000	2000US-02363.68
PR	02-OCT-2000	2000US-02367.02
PR	02-OCT-2000	2000US-02370.37
PR	02-OCT-2000	2000US-02370.38
PR	02-OCT-2000	2000US-02370.40
PR	13-OCT-2000	2000US-02393.37
PR	13-OCT-2000	2000US-02393.37
PR	20-OCT-2000	2000US-02418.62
PR	20-OCT-2000	2000US-02418.65
PR	20-OCT-2000	2000US-02418.81
PR	08-NOV-2000	2000US-02464.74
PR	08-NOV-2000	2000US-02464.76
PR	08-NOV-2000	2000US-02464.76
PR	08-NOV-2000	2000US-02464.77
PR	08-NOV-2000	2000US-02464.78
PR	08-NOV-2000	2000US-02465.23
PR	08-NOV-2000	2000US-02465.24
PR	08-NOV-2000	2000US-02465.25
PR	08-NOV-2000	2000US-02465.26
PR	08-NOV-2000	2000US-02465.27
PR	17-NOV-2000	2000US-02492.07
PR	17-NOV-2000	2000US-02492.08
PR	17-NOV-2000	2000US-02492.09
PR	17-NOV-2000	2000US-02492.10
PR	17-NOV-2000	2000US-02492.11
PR	17-NOV-2000	2000US-02492.12
PR	17-NOV-2000	2000US-02492.14
PR	17-NOV-2000	2000US-02492.15
PR	17-NOV-2000	2000US-02492.16
PR	17-NOV-2000	2000US-02492.17
PR	17-NOV-2000	2000US-02492.18
PR	17-NOV-2000	2000US-02492.19
PR	17-NOV-2000	2000US-02492.20
PR	17-NOV-2000	2000US-02492.21
PR	17-NOV-2000	2000US-02492.22
PR	17-NOV-2000	2000US-02492.23
PR	17-NOV-2000	2000US-02492.24
PR	17-NOV-2000	2000US-02492.25
PR	17-NOV-2000	2000US-02492.26
PR	17-NOV-2000	2000US-02492.27
PR	17-NOV-2000	2000US-02492.28
PR	17-NOV-2000	2000US-02492.29
PR	17-NOV-2000	2000US-02492.30
PR	17-NOV-2000	2000US-02492.31
PR	17-NOV-2000	2000US-02492.32
PR	17-NOV-2000	2000US-02492.33
PR	17-NOV-2000	2000US-02492.34
PR	17-NOV-2000	2000US-02492.35
PR	17-NOV-2000	2000US-02492.36
PR	17-NOV-2000	2000US-02492.37
PR	17-NOV-2000	2000US-02492.38
PR	17-NOV-2000	2000US-02492.39
PR	17-NOV-2000	2000US-02492.40
PR	17-NOV-2000	2000US-02492.41
PR	17-NOV-2000	2000US-02492.42
PR	17-NOV-2000	2000US-02492.43
PR	17-NOV-2000	2000US-02492.44
PR	17-NOV-2000	2000US-02492.45
PR	17-NOV-2000	2000US-02492.46
PR	17-NOV-2000	2000US-02492.47
PR	17-NOV-2000	2000US-02492.48
PR	17-NOV-2000	2000US-02492.49
PR	17-NOV-2000	2000US-02492.50
PR	17-NOV-2000	2000US-02492.51
PR	17-NOV-2000	2000US-02492.52
PR	17-NOV-2000	2000US-02492.53
PR	17-NOV-2000	2000US-02492.54
PR	17-NOV-2000	2000US-02492.55
PR	17-NOV-2000	2000US-02492.56
PR	17-NOV-2000	2000US-02492.57
PR	17-NOV-2000	2000US-02492.58
PR	17-NOV-2000	2000US-02492.59
PR	17-NOV-2000	2000US-02492.60
PR	17-NOV-2000	2000US-02492.61
PR	17-NOV-2000	2000US-02492.62
PR	17-NOV-2000	2000US-02492.63
PR	17-NOV-2000	2000US-02492.64
PR	17-NOV-2000	2000US-02492.65
PR	17-NOV-2000	20

PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251859.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251989.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0253678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
XX	WPI, 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
XX		
XX	Disclosure; SEQ ID NO 29573; 3071bp + Sequence Listing; English.	
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patient's own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK67694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
XX	Sequence 481 BP; 140 A; 103 C; 79 G; 159 T; 0 other:	
Query Match	3.4%; Score 75.8; DB 22; Length 481;	
Best Local Similarity	68.9%; Pred. No. 8e-12;	
Matches 104; Conservative	0; Mismatches 47; Indels	0; Gaps
0Y	12cttcctgtaagaaccagacaggaactatttagctgtgtgcatatgctcagct	971
11		
Db	282 tttttttgtaaggccagacaggaatatttttagctgtgtgcatatgctcagct	341
0Y	972 ccaacactatctctctgtgctgtgagcaggaagaattagcaacaatgtaacaaa	1031
11		
Db	342 tgcacatgtctcagtttgcctctgtagacctaagaagccgtgatgataccta	401
0Y	1032 catatgtgaccccatgaaactatttatt	1062
11		
Db	402 cgggtgtgactgtgttccaaaaaatatttatt	432
RESULT	7	
ABA04430		
ID	ABA04430 standard; cDNA; 1837 BP.	
XX		
AC	ABA04430;	
XX		
DT	11-MAR-2002 (first entry)	
XX		
DE	Human PP1345 protein encoding cDNA SEQ ID NO:19/21.	
XX		
KW	Human; PP1345; cancer suppression; ss.	
XX		

```

OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 174..566
XX FT /tag="a
XX FT /product="PPI345"
XX
XX CN1313315-A.
XX
XX PD 19-SEP-2001.
XX
XX PF 13-MAR-2000; 2000CN-0111989.
XX
XX PR 13-MAR-2000; 2000CN-0111989.
XX
XX PA (SHAN-) SHANGHAI INST ONCOLOGY.
XX
XX PI Gu J, Yang S;
XX
XX DR WPI: 2002-042193/06.
XX
XX DR P-PSDB; ABB04713.
XX
XX PT New human protein able to suppress growth of cancer cells and its
XX encoding polynucleotide sequence.
XX
XX PS Claim 5; Page 33 (Disclosure); 42pp; Chinese.
XX
XX CC The present sequence encodes human PPI345 protein, which has cancer-
XX suppressing activity. The present invention also describes a method
XX for the preparation of the protein by recombination, and the application
XX of the protein in treating diseases such as cancer.
XX
XX SQ Sequence 1837 BP; 478 A; 471 C; 407 G; 481 T; 0 other;

Query Match 3.3%; Score 73; DB 24; Length 1837;
Best Local Similarity 64.5%; Pred. No. 1.2e-10;
Matches 109; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 918 tgaagaacacagacaggaactatttagctctgtgacatagtctcgtcacac 977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1639 tttaagaagacagatagtaatttttaagctctgtgacacatgtctcgtcacac 1698

OY 978 tactcatctctgctctgtagcagaagaacttagcaacaatattgcaacaacatag 1037
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1699 tattcagctctgtctatgaatacaaaagcgtgtgacagatgttaaatgaatgatg 1758

OY 1038 tgaccacagaaacttatttatttgatagtaggaacctgaaataa 1086
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1759 tggctgtgtccataaacttatttcccaaaaaa 1807

RESULT 8
AAH98619/C
ID AAH98619 standard; cDNA; 628 BP.
XX
XX AC AAH98619;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Human EST-derived coding sequence SEQ ID NO: 476.
XX
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX diagnostics; forensic test; gene mapping; genetic disorder;
XX biodiversity; gene therapy; nutrition; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200154477-A2.
XX
XX PD 02-AUG-2001.
XX

```

```

PF 25-JAN-2001; 2001WO-US02687.
XX
XX PR 25-JAN-2000; 2000US-0491404.
XX
XX PR 17-JUL-2000; 2000US-0617246.
XX
XX PR 03-AUG-2000; 2000US-0611451.
XX
XX PR 15-SEP-2000; 2000US-0663870.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YF, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX Cao Y, Dimaheac RA, Zhang J, Werhman T;
XX
XX DR WPI: 2001-476164/51.
XX
XX DR P-PSDB; AAM23960.
XX
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use.
XX
XX PS Claim 1; Page 514; 1275pp; English.
XX
XX CC The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX from the organism of interest. They can be used in diagnostics,
XX forensics, gene mapping, identification of mutations, to assess
XX biodiversity and for nutritional purposes. The present sequence is a cDNA
XX of the invention.
XX
XX SQ Sequence 628 BP; 190 A; 122 C; 131 G; 185 T; 0 other;

Query Match 3.2%; Score 71.6; DB 22; Length 628;
Best Local Similarity 60.9%; Pred. No. 1.6e-10;
Matches 151; Conservative 0; Mismatches 94; Indels 3; Gaps 2;

OY 894 gacccaggttgggcaactcttctcgttaagaacacaggaactatttagctcgt 953
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 452 GACCAGGGTTCAGCAAGCTTTTCTGTAAAGCGCAAAATGTAATTTAGCTTTGT 393

OY 954 gtccatattggtcagtcagcaacactcactcctcgtctgtagcagaagaacatag 1013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 392 GGACCATATGCTC--TGCTACTACTCAATTCGCCATTTGAAAGTGAAGACACAT 335

OY 1014 caacatatgtcacaacaacatattgaccccatgaaacttatttattgatacgga 1073
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 GTA-AATGATGTGACGTGGCTGTTTATTTGTGACACACTGAAATTTAAATTCATAATT 276

OY 1074 aacctgaataatattgcttcttcttgaatttttcccaatcaataaagaagtaaaact 1113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 275 TTCACATAAAATATATTCCTTTTGATTTTTCGCCAACAATTAATAATTTAAAGCC 216

OY 1134 acctctag 1141
    |||||
DB 215 ATGCTTAG 208

RESULT 9
AAA41353/C
ID AAA41353 standard; cDNA; 286 BP.
XX
XX AC AAA41353;
XX
XX DT 21-AUG-2000 (first entry)
XX
XX DE Human secreted expressed sequence tag SEQ ID NO:93.
XX
XX KW Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
XX expressed sequence tag; EST; probe; chemotactic; analgesic; haemostatic;
XX immunomodulatory; haematopoietic; chemokinetic; antiinflammatory;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
XX antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX

```

KW	cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW	vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW	central nervous system disorder; Alzheimer's disease; stroke;
KW	Parkinson's disease; Huntington's disease; coagulation disorder;
KW	haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW	tumour; infection; depression; psoriasis; ss.
XX	
OS	Homo sapiens.
PN	WO200021990-A1.
XX	
PD	20-APR-2000.
XX	
PF	15-OCT-1999; 99NO-US24205.
PR	15-OCT-1998; 98US-0104435.
PA	(GENY ) GENETICS INST INC.
PI	Jacobs K, MCCoy JM, LaValle ER, Collins-Racle LA, Evans C;
PI	Meredys D, Treacy M;
DR	WPt; 2000-317937/27.
XX	
PT	Isolated polynucleotides, and encoded proteins, comprising secreted
PT	expressed sequence tags (SESTs), useful for treating various disorders
PT	such as autoimmune, infectious, and central nervous system disorders -
PS	Claim 1; Page 198, 618pp; English.
XX	
CC	AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC	sequence tags (SESTs), isolated from human, mouse, xenopus and rat
CC	tissue sources. The SESTs can have a range of activities depending on
CC	the tissues they were isolated from. The activities include:
CC	chemotactic; proliferative; immunomodulatory; haematopoietic;
CC	chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC	cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC	antiasthmatic; vulnerrary; anticancer; osteopathic; neuroprotective;
CC	neurologic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC	anticonvulsant; and antidepressant. The SESTs can be used for gene
CC	therapy and in vaccines. The SESTs are useful as probes for the
CC	identification and isolation of full-length cDNAs and genomic DNA
CC	molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC	are useful in assays for determining biological activity and raising
CC	antibodies. They may be useful for treatment of autoimmune disorders
CC	(multiple sclerosis, insulin dependent diabetes), allergic conditions
CC	(asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC	osteoporosis, osteoarthritis, central nervous system disorders
CC	(Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC	disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC	disease), tumours, bacterial, fungal or viral infections, depression and
CC	psoriasis. AAA43420 to AAA43425 represent linker variants which are given
XX	in the exemplification of the present invention.
XX	
XX	Sequence 286 BP; 88 A; 63 C; 50 G; 85 T; 0 other;
XX	
Query Match	3.2%; Score 71.4; DB 21; Length 286;
Best Local Similarity	72.1%; Pred. No.1,2e-10;
Matches 93; Conservative	0; Mismatches 36; Indels 0; Gaps 0
Db	
899	aggcttgcaaacactctctcgttaagaaccagacaactatttaggcctgvgtc 958
149	AGGGCAACATTAACCTTTTTCGTGTAACAGTCAACAGATAATTATTTTAGCCTCGTGTC 90
Oy	atacgagctatgcacaaactcatcctctgacctgtgatgccgaagaacattggcaaca 1018
Db	89 ATATAGCTCTCTGTCGCAACTACCAACTCTGCTGTTCTTGTTGTAAGAACAATAGCA 30
Oy	1019 atagtcaa 1027

[illegible]

DB 224 ACATTGAATTGAA 210

## RESULT 11

AAH8703  
ID AAH8703 standard; DNA; 123219 BP.  
XX  
AC AAH8703;  
XX  
XX 26-FEB-2002 (first entry)  
XX  
DE Human DNA sequence SEQ ID 543.  
XX  
XX Single nucleotide polymorphism; SNP; diallelic marker; human;  
XX central nervous system disorder; CNS; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200151659-A2.  
XX  
XX 19-JUL-2001.  
XX  
XX 11-JAN-2001; 2001WO-1B00116.  
XX  
XX 13-JAN-2000; 2000US-0175854.  
XX  
XX (GEST ) GENSET.  
XX  
XX Chu T, Blumenfeld M, Cohen D;  
XX  
XX WPI; 2001-483085/52.  
XX  
XX Isolated polynucleotides, useful for genotyping nucleic acids for  
XX diallelic markers for the diagnosis of depression, comprises central  
XX nervous system disorder related diallelic marker -  
XX  
XX Disclosure: Page 439-472; 51pp; English.  
XX  
XX The present invention relates to diallelic markers derived from human  
XX genes involved in central nervous system (CNS) disorders (see  
XX AAH88161-AAH88702). The markers have a single nucleotide polymorphism  
XX (SNP) and are useful in determining the genetic predisposition of  
XX individuals to CNS disorders, by identifying the nucleotides at a set of  
XX genetic markers in a biological sample, where the markers comprise at  
XX least one CNS disorder related marker. The present sequence was used  
XX to illustrate the invention.  
XX  
XX Sequence 123219 BP; 27921 A; 32327 C; 32369 G; 30437 T; 165 other;

Query Match 3.1%; Score 70; DB 23; Length 123219;  
Best Local Similarity 61.5%; Pred. No. 9.7e-09;  
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 895 acccaggtggggaacctctctgtgaagaacacagaggaactatttagctctgtg 954  
DB 17375 agcagggggtcgaacactcttcttgaaaggagatagtaacgtgcataaggcttctg 17434  
QY 955 tgcacatagctcagtcacaaactactctctctgtcctctgtgacagaagaactaac 1014  
DB 17435 ggcacatagctcagtcacaaactactctctctgtcctctgtgacagaagaactaac 17494  
QY 1015 aacacatagtcacaaacatactgaccccaagaaacttatttattgatacagaa 1074  
DB 17495 gacacatagtcacaaaggccgtgtctgcatgccaataaacttattataaacaac 17554  
QY 1075 ac 1076  
DB 17555 ac 17556

## RESULT 12

AAK3517/C  
ID AAK3517 standard; DNA; 50000 BP.

## AAK3517;

DT 23-JUN-1999 (first entry)

DE Human kidney aminopeptidase P genomic DNA fragment 1.

AAminopeptidase; human; Amp; gene therapy; treatment; Amp-deficiency;  
XX prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;  
XX arterial stenosis; industrial protein feed; malabsorption syndrome;  
XX proteinaceous waste degradation; additive; immunohistochemistry; ss.  
XX

OS Homo sapiens.

PN WO9911799-A2.

PD 11-MAR-1999.

PE 02-SEP-1998; 98WO-US18426.

PR 02-SEP-1997; 97US-0057854.

PA (MEDT-) MEDICAL COLLEGE GEORGIA RES INST.

PI Ryan JW, Sprinkle JJC, Venema RC;

DR WPI; 1999-205193/17.

PT Nucleic acid encoding human aminopeptidase P

PS Claim 13; Page 80-109; 201pp; English.

XX This invention describes the isolation of a novel human aminopeptidase P  
XX (Amp). This protein is used to produce recombinant Amp and can be used  
XX for gene therapy for treating Amp-deficiency conditions. Its fragments  
XX are used as primers and probes to identify patients with homozygous and  
XX heterozygous Amp deficiency, including prenatal diagnosis (patients  
XX defective in Amp are at risk of developing angioedema if treated with  
XX angiotensin-converting enzyme inhibitors), also as antisense inhibitors  
XX in cases of excessive Amp expression. The product of the invention is  
XX also used to identify Amp-expressing sequences in other animals and to  
XX generate transgenic animals, and comparisons of genomic sequences are  
XX used to detect mutations. Amp inhibitors are potentially useful as  
XX antihypertensive agents and to prevent or treat arterial (re)stenosis  
XX or atherosclerosis. The structure of Amp is used to design synthetic  
XX substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal  
XX imido bonds, can be used to degrade industrial protein feeds to free  
XX amino acids, to degrade proteinaceous wastes, as additives in enzyme  
XX formulations used to treat malabsorption syndrome and for studying its  
XX biological role. Antibodies against Amp are used in immunohistochemical  
XX methods to study Amp distribution.  
XX

XX Sequence 50000 BP; 13187 A; 12125 C; 11842 G; 12846 T; 0 other;

Query Match 3.1%; Score 69.8; DB 20; Length 50000;  
Best Local Similarity 72.7%; Pred. No. 6.7e-09;  
Matches 104; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 889 gctctgaccaggttgggcaacactctctgtgaagaacacagaggaactatttaggc 948  
DB 8967 gctctgacatcaaaagctttgccaaa--ctttctgtgaaagcgctacagatgtaagatattttacac 8910  
QY 949 tctgtgtccatatactgctcagtcacaaactactcctctgctctgtagcaagaagca 1008  
DB 8909 tcttcgacacacatcagctgctctgtacacacacactcactcgtctgtgacagcccaaacca 8850  
QY 1009 attagcaacaatctgcaacaaa 1031  
DB 8849 agtatagacaaatattataaacaaa 8827

XX	RESULT_13
XX	AHA99747
XX	ID AHA99747 standard; cDNA; 8029 BP.
AC	
XX	AHA99747:
XX	
DT	16-OCT-2001 (first entry)
XX	
DE	Human protein encoding cDNA sequence SEQ ID NO:582.
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KM	antiflammatory; antirheumatic; antiarthritic; immunosuppressive;
KM	antibacterial; endocrine; cardiac; central nervous system; virucide;
KM	anti-HIV; fungicide; antimutagen; cardiovascular; antihaemic; anaemia;
KM	angioregiant; haemostatic; vulnery; anticancer; osteopathic; eczema;
KM	dematological; allergic; antistatic; antidiabetic; cytostatic;
KM	neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KM	immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KM	antiapophytic; rheumatoid arthritis; septic shock; pancreatitis;
KM	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KM	genetic disease; hematopoietic disorder; platelet disorder; asthma;
KM	thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KM	allergic rhinitis; diabetes; multiple sclerosis; depression;
KM	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KV	neurological disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153455-A2.
XX	
PD	26-JUL-2001.
XX	
PE	22-DEC-2000; 2000WC-US35017.
XX	
PR	23-DEC-1999; 99US-0471275.
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
XX	
PA	(HSE-) HYSEQ INC.
XX	
P1	Tang YT, Liu C, Drmanac RT;
DR	MP1: 2001-457603/49.
DR	P-PSDB: AAM25806.
PT	
PT	Isolated human polynucleotides encoding polypeptides, useful for the
PT	treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX	
PS	Claim 1; Page 622-624; 1217pp; English.
CC	
CC	AHA99166 to AHA99904 encode the human proteins given in AAM25225 to
CC	AAM25363. The proteins can have activities based on the tissues and
CC	cells they are expressed in, such as: antiflammatory; antirheumatic;
CC	antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC	central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC	cardiovascular; antihaemic; angioregiant; haemostatic; vulnery;
CC	anticancer; osteopathic; dermatological; antiallergic; antiastmatic;
CC	antidiabetic; cytotatic; neuroprotective; antidepressant; nootropic;
CC	antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC	encoding them can be used in gene therapy, antisense therapy and vaccine
CC	production. The proteins and polynucleotides are useful for screening for
CC	agonists or antagonists of a protein and for the treatment and diagnosis
CC	of disorders associated with the activity of a protein e.g. Inflammation,
CC	rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC	neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC	infections, autoimmunity, genetic diseases, hematopoietic disorders,
CC	anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC	osteoporosis, severe combined immunodeficiency, eczema, allergic
CC	rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC	Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC	neurological disorders.
XX	

Seq	Sequence	8029 BP	2064 A	1942 C	1905 G	2118 T	0 other:
Query Match	3.1%;	Score 69.4;	DB 22;	Length 8029;			
Best Local Similarity	57.2%;	Pred. NO. 3.1e-09;					
Matches 202;	Conservative 0;	Mismatches 121;	Indels 30;	Gaps 3			
OY	888 gccctgagccaggttggcgcaactcttcctgtaagaagaccagagactatcttaggc	948					
DB	682 gctctgattcagagctgcggcaactcttctttaagagacaagtagtaattctttaggc	741					
OY	949 tctgtgtgccatgtgtctcagctcacactactatctctctgcctctgtagcac-----	1001					
DB	742 ttgtcagcgtactaacgtccctctgccaaactatctaacctctgcgaagaagccagtcaccga	801					
OY	1002 -----gaaagcaattagcaacaatalatgtcaacaacacatatgtgaccccatga	1048					
DB	802 caalacataaactgatatggtgtgtgctatgctctataacatttgtttatgacactga	861					
OY	1049 aaactcttctatcatgatacgcgaaacctgaaataatgctcttcttttatcttttc	1107					
DB	862 aactcgaattccatgaattcttcacaggtgcacgaatatcttctctttgtatcttcg	921					
OY	1108 cccaatcattaaaaacgtaaaaactactctttaggtcgcgaaggttaagccatctcagct	1167					
DB	922 cccaatcatttgaacaacataaccatctcttagctcacag-----ctgataaa	972					
OY	1168 tagcagtgcgacgctcgtgatttgccttgactacagttggccaatccctgatt	1220					
DB	973 aacaggtgcagcgccaggtttgattcgttgagccatagtttgccaacctgatt	1025					
RESULT 14							
AA599306/C							
ID	AA599306 standard; DNA; 46765 BP.						
XX	AA599306;						
AC							
XX							
DT	12-MAR-2002 (first entry)						
XX							
XX							
DE	DNA encoding Aldehyde dehydrogenase 5 family, member A1 (ALDH5A1).						
KW	Aldehyde dehydrogenase 5 family member A1; ALDH5A1;						
KW	succinate-semialdehyde dehydrogenase; gene therapy; probe;						
KW	antisense technology; allele specific oligonucleotide; ASO;						
KW	4-hydroxybutyric aciduria; metabolic disease; transgenic animal;						
KW	chromosome 6p22; ds.						
XX							
OS	Homo sapiens.						
XX							
XX							
FT	Key	Location/Qualifiers					
FT	Variation	replace(3749..T)					
FT		/*tag= a					
FT	Variation	/*standard_name= replace(3800..G)					
FT		/*tag= b					
FT	Variation	/*standard_name= replace(3829..T)					
FT		/*tag= c					
FT	Variation	/*standard_name= replace(3875..G)					
FT		/*tag= d					
FT	Variation	/*standard_name= replace(3876..A)					
FT		/*tag= e					
FT	Variation	/*standard_name= replace(3941..T)					
FT		/*tag= f					
FT	Variation	/*standard_name= replace(4050..4403)					
FT		/*tag= g					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= h					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= i					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= j					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= k					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= l					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= m					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= n					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= o					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= p					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= q					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= r					
FT	Variation	/*standard_name= replace(4050..42765)					
FT		/*tag= s					

FT		/tag= h	
FT		/product= "ALDH5A1"	
FT		/note= "Aldehyde dehydrogenase 5 family member A1. Specifically claimed in claim 25"	
FT	variation	replace(4059,G)	
FT		/tag= i	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(4111,A)	
FT		/tag= j	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(4155,C)	
FT		/tag= k	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(4179,G)	
FT		/tag= l	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	Intron	4404..11575	
FT		/tag= m	
FT		/number= 1	
FT	variation	replace(4430,G)	
FT		/tag= n	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	exon	11576..11659	
FT		/tag= o	
FT		/number= 2	
FT	Intron	11660..12315	
FT		/tag= p	
FT		/number= 2	
FT	exon	12316..12486	
FT		/tag= q	
FT		/number= 3	
FT	variation	replace(12415,T)	
FT		/tag= r	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(12422,T)	
FT		/tag= s	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	Intron	12487..13921	
FT		/tag= t	
FT		/number= 3	
FT	variation	replace(12559,T)	
FT		/tag= u	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(13855,G)	
FT		/tag= v	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(13917,T)	
FT		/tag= w	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	exon	13922..14038	
FT		/tag= x	
FT		/number= 4	
FT	variation	replace(14021,T)	
FT		/tag= y	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	Intron	14039..24219	
FT		/tag= z	
FT		/number= 4	
FT	variation	replace(14046,A)	
FT		/tag= aa	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	variation	replace(14179,A)	
FT		/tag= ab	
FT		/standard_name=	"Single nucleotide polymorphism"
FT	exon	24220..24363	
FT		/tag= ac	
FT		/number= 5	
FT	Intron	24364..29453	
FT		/tag= ad	
FT		/number= 5	
FT	Intron	24598..31819	
FT		/tag= ae	
FT		/number= 6	

FT	variation	replace(29397,T)
FT		/*tag= at
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	29454...29597
FT		/*tag= ag
FT		/number= 6
FT	variation	replace(29469,A)
FT		/*tag= ah
FT	exon	/standard_name= "Single nucleotide polymorphism"
FT		31820...31978
FT		/*tag= ai
FT		/number= 7
FT	intron	31979...37049
FT		/*tag= aj
FT		/number= 7
FT	exon	37050...37219
FT		/*tag= ak
FT		/number= 8
FT	intron	37220...41171
FT		/*tag= al
FT		/number= 8
FT	exon	41172...41230
FT		/*tag= am
FT		/number= 9
FT	variation	replace(41217,C)
FT		/*tag= an
FT		/standard_name= "Single nucleotide polymorphism"
FT	intron	41231...42559
FT		/*tag= ao
FT		/number= 9
FT	variation	replace(41237,A)
FT		/*tag= ap
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(41403,C)
FT		/*tag= aq
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(42380,T)
FT		/*tag= ar
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	42560...42765
FT		/*tag= as
FT		/number= 10
PN		WO200190119-A2.
XX		
PD		29-NOV-2001.
XX		
PF		21-MAY-2001; 2001WO-US16558.
XX		
PR		19-MAY-2000; 2000US-205849P.
XX		
PA		(GENA-) GENAISSANCE PHARM INC.
XX		
PI		Kliem SE, Koshy B, Tanguay DA;
DR		WPI: 2002-089912/12.
DR		P-PSDB: AAU73594.
XX		
PT		New genetic variants of human aldehyde dehydrogenase 5 family, member
PT		A1, ALDH5A1 gene for treating metabolic diseases and for expressing
PT		ALDH5A1 protein useful in identifying drugs to treat 4-hydroxybutyric
PT		aciduria -
XX		
PS		Disclosure: Fig 1; 151pp; English.
XX		
CC		The invention describes an isolated polynucleotide comprising a
CC		nucleotide sequence which is a polymorphic variant of a reference
CC		sequence for the aldehyde dehydrogenase 5 family, member A1
CC		(succinate-semialdehyde dehydrogenase) (ALDH5A1) gene or its fragment.
CC		The polypeptide is useful for screening for drugs targeting it by
CC		contacting the ALDH5A1 polymorphic variant with a candidate agent and
CC		assaying for binding activity. The polypeptide and haplotypes are useful
CC		for identifying an association between a trait such as a clinical





PR	20-OCT-2000	2000US-0241721
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241826
PR	01-NOV-2000	2000US-0244617
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249264
PR	17-NOV-2000	2000US-0249265
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249299
PR	17-NOV-2000	2000US-0249300
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0251030
PR	05-DEC-2000	2000US-0251088
PR	05-DEC-2000	2000US-0251679
PR	06-DEC-2000	2000US-0251479
PR	08-DEC-2000	2000US-0251856
PR	08-DEC-2000	2000US-0251868
PR	08-DEC-2000	2000US-0251869
PR	08-DEC-2000	2000US-0251989
PR	08-DEC-2000	2000US-0251990
PR	11-DEC-2000	2000US-0254097
PR	05-JAN-2001	2001US-0259678

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PT Rosen CA, Barash SC, Ruben SM,  
XX  
DR WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 19625; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (1)  
CC amino acid sequences given in AAM62170 to AAM91921. (1) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (1)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 47319 BP; 13708 A; 8453 C; 9720 G; 15438 T; 0 other;

Query Match	3.1%	Score 69.4	DB 22	Length 47319
Best Local Similarity	66.2%	Pred. No. 8.5e-09		
Matches 100	Conservative	0	Mismatches 51	Indels 0
Qy 881	ttcccatgcctgagccccaggttcgggcaaacctccctgtaagaaccagagaacta	940		
Db 18549	TTTCTAGCTTAATCAGAGATTGGCAACTTTTCTGTCAGAGCCAGGCATTAATTA	18490		
Qy 941	ttttaggtctctgltggtccatgatgltctcagtcacacactatcatctctgctgtagca	1000		
Db 18489	TTTTAGGCTTTGTTAAGCCACACAGCTCTTACGGCAACCACTCACTGTCATATAGCC	18430		
Qy 1001	cgaagcaaatatgacacacatctgcacaaa	1031		
Db 18429	CCAAAGCTGTACAGATGATATTAATCAA	18399		

Search completed: August 16, 2002, 05:43:35  
Job time: 11519 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:40:06 ; Search time 3499.63 Seconds  
(without alignments)  
8816.372 Million cell updates/sec

Title: US-09-515-369B-1  
Perfect score: 2286  
Sequence: 1 taatacgaactacatacaggg.....tgacttccagcgcgtggagc 2286

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthbm:\*  
3: em\_estln:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pla:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	3.5	480	10 B1262402	B1262402 602953836
2	80.6	3.5	354	10 T58770	T58770 Yb80h02.s1
3	80.4	3.5	390	10 L44398	L44398 HUMEST1E3 H
4	78.4	3.4	521	12 A0695956	A0695956 HS-2160_A
5	78.2	3.4	435	9 AM183569	AM183569 xj76h04.x
6	78.2	3.4	459	9 A1125442	A1125442 qd89b02.x
7	78.2	3.4	628	12 A0378792	A0378792 RPI11-15
8	77.4	3.4	393	9 A1128823	A1128823 qe94d06.s
9	77.4	3.4	411	10 W76494	W76494 z67d03.r1
10	77.4	3.4	680	12 AG154552	AG154552 Pan trogl
11	77.2	3.4	400	9 A1926274	A1926274 w043g06.x
12	77.2	3.4	322	9 A1240516	A1240516 qh12h02.x
13	77.2	3.4	386	9 AA443938	AA443938 zv51c03.s
14	77.2	3.4	393	9 AA444117	AA444117 zv51c03.r
15	77.2	3.4	417	9 A1208768	A1208768 q938c10.x
16	76.8	3.4	509	9 B1061153	B1061153 t13-pt011
17	76.6	3.4	327	12 B98848	B98848 CIT-HSP-228

18	76.2	3.3	499	12 A0277351	A0277351 CITBI-E1-
19	76.2	3.3	644	12 A0039307	A0039307 CIT-HSP-2
20	76.2	3.3	5529	12 AF101969	AF101969 AF101969
21	76.2	3.3	573	12 A0633901	A0633901 RPCI-11-4
22	75.6	3.3	530	12 A0508112	A0508112 RPCI-11-2
23	75.6	3.3	628	12 A0020407	A0020407 CIT-HSP-2
24	75.6	3.3	799	10 BG535645	BG535645 60256353
25	75.2	3.3	488	9 AV604853	AV604853 AV604853
26	74.6	3.3	348	12 A0035111	A0035111 CIT-HSP-2
27	73.6	3.2	640	12 AG093736	AG093736 Pan trogl
28	73.2	3.2	381	9 AM574882	AM574882 UT-HF-BN0
29	73.2	3.2	512	12 A0768179	A0768179 HS-3160_B
30	73.2	3.2	519	12 A0029411	A0029411 RPCI11-39
31	72.8	3.2	302	9 A1559884	A1559884 h87a12.x
32	72.8	3.2	427	12 A0815705	A0815705 HS-5394_A
33	72.2	3.2	312	10 BG230513	BG230513 n014406-
34	72.2	3.1	416	9 AM835797	AM835797 OVA-LT001
35	72.2	3.1	764	10 B1915216	B1915216 603184745
36	71.8	3.1	2032	11 AF161549	AF161549 Homo sapi
37	71.4	3.1	794	10 BF983586	BF983586 602306575
38	71.2	3.1	327	9 A1344682	A1344682 qp09h04.x
39	71.2	3.1	363	9 A1358986	A1358986 qy19h02.x
40	71.2	3.1	372	9 A1370636	A1370636 ta38g12.x
41	71.2	3.1	380	9 A1244892	A1244892 qj98f05.x
42	71.2	3.1	380	9 A1298789	A1298789 qm91c02.x
43	71.2	3.1	388	9 A1291781	A1291781 qm75b05.x
44	71.2	3.1	393	9 A1214952	A1214952 qm31d12.x
45	71.2	3.1	394	9 A1439755	A1439755 t161a07.x

## ALIGNMENTS

RESULT 1  
LOCUS B1262402 480 bp mRNA linear EST 17-JUL-2001  
DEFINITION 602953836T1 NIH\_MGC\_99 Homo sapiens CDNA clone IMAGE:5088033 3',  
mRNA sequence.  
ACCESSION B1262402.1 GI:14822582  
VERSION B1262402.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (Bases 1 to 480)  
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA Library  
Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1846 row: 9 column: 10  
High quality sequence start: 27  
High quality sequence stop: 480.  
Location/Qualifiers  
1. 480  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5088033"  
/clone\_lib="NIH\_MGC\_99"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph. Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GCCACGAC(C). Size-selected >500bp for average insert size





A1125442 459 bp mRNA linear EST 28-OCT-1998  
 DEFINITION g089b02.x1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:1736619  
 ACCESSION A1125442  
 VERSION A1125442.1 GI:3593956  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 459)  
 NCBI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-rr@mail.nih.gov](mailto:cgaps-rr@mail.nih.gov)  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
 , Ph.D.

FEATURES  
 source  
 Location: Qualifiers  
 1..459  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1736619"  
 /clone\_1db="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: p7773D-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - 0149(dT) primer [5].  
 TGTTCACATCTGAAAGTGGGCGCGCCCAATTTTCTTTT 3'}.  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p7773 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 159 a 88 c 61 g 151 t  
 ORIGIN

Query Match 3.4% Score 78.2; DB 9; Length 459;  
 Best Local Similarity 59.1%; Pred. No. 3.6e-09;  
 Matches 204; Conservative 0; Mismatches 113; Indels 28; Gaps 3;

Oy 901 gctgggaacactctccgtgaagaacagagaaacttttaggtctgtgtgcat 960  
 Db 105 GGTGACGAAACTTTCTGTAAGGGCCATACATACATATTCAGTCTGCGAGACAT 164  
 Oy 961 atggtctcagtcacacactctctgtgacctgtgacagaaagaattagcaaacat 1020  
 Db 165 TTGGTCTTTGCGAAACTCTGTAAGTCTGTAACATACAAAGCTGCCACAGACAT 224  
 Oy 1021 atgtcaacaacatgtgaccccaatgaacattattat--tatgatacggaaacc 1077  
 Db 225 ATGTAACAAATGAGGGTGAAGTAAATGATTAATTAAGACACTGAATTTGAACT 284  
 Oy 1078 tgaataatgtctctct-----tttatctttcccaatcatataaa 1121  
 Db 285 TCAATCATTTTCATCTCCCAAAACATTTACTTTTGACTATTTTCAACCACTTAAA 344  
 Oy 1122 aacgtataaactactcttaggtcgaaggttaagccattcagcttagcagtcaggc 1181  
 Db 345 AAAGTAAACCACTTCTTAGTGACAGAGCTGAGAAAACAT-----AGTCATAGGC 395

Oy 1182 tgaattgtgcttctgtgacctacagttggtccaaatccctgattcccaa 1226  
 Db 396 TTAATCTGGCCACACAGCATATGTTGCAATGCGCTGTACACTA 440

RESULT 7  
 LOCUS A0378792/c  
 DEFINITION RC111-151B24.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-151B24,  
 DNA sequence.  
 ACCESSION A0378792  
 VERSION A0378792.1 GI:4349815  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 628)  
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter  
 , J.C.  
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other GSSs: RC111-151B24.TJ  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0208  
 Fax: 301 838 0200  
 Email: [hbeetlgr.org](mailto:hbeetlgr.org)

FEATURES  
 source  
 Location/Qualifiers  
 1..628  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-151B24"  
 /clone\_1db="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"

BASE COUNT 173 a 133 c 136 g 185 t 1 others  
 ORIGIN

Query Match 3.4% Score 78.2; DB 12; Length 628;  
 Best Local Similarity 72.1%; Pred. No. 4e-09;  
 Matches 101; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Oy 892 tgaacccaggttggcaaacctctctgtgaagaacagagaaactatttaggtctc 951  
 Db 572 TAGGCCAGGGGTTGGCTACTTTCTGTAAGACCCAGATGGAATATTTTAATTNT 513  
 Oy 952 gtgtgcataatgttgcagtcacactactctctgtgacctgtgacagaaagaact 1011  
 Db 512 GTGGCCCTTAAAGCTCTGTAACACCACTAATTTCTGAGCTAGCATGAAGAGGCC 453  
 Oy 1012 agcaacaatgtcacaana 1031  
 Db 452 ACAGACAGTAAGTAACAAA 433

RESULT 8  
 A1128823



LOCUS	A1128823	393 bp	mRNA	linear	EST 11-SEP-1998
DEFINITION	g994d06.s1 Soares_fetal_heart.NBHH19W Homo sapiens cDNA clone IMAGE:1694411 3', mRNA sequence.				
ACCESSION	A1128823				
VERSION	A1128823.1 GI:3597337				
KEYWORDS	EST.				
ORGANISM	human.				
SOURCE	Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 393) NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/cgi/gap">http://www.ncbi.nlm.nih.gov/cgi/gap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: crabs-f@mail.nih.gov This clone is available royalty-free through LMTL; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information. Seq primer: -40ml3 fwd. ET from Amer sham.				
FEATURES	Location/Qualifiers				
source	1..393				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone-"IMAGE:1694411"				
	/clone_id="Soares_fetal_heart_NBHH19W"				
	/sex="unknown"				
	/dev_stage="19 weeks"				
	/_lab_host="DH10B (ampicillin resistant)"				
	/_note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dt) primer [5' TGTTACCATCTGTGAAGTCGGAGCGCGGCATCTTTTTTTTTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into The Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Felina Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHH19W."				
BASE COUNT	140 a	70 c	53 g	130 t	
ORIGIN					
Query Match	3.4%; Score 77.4; DB 9; Length 393;				
Best Local Similarity	60.2%; Pred. No. 5.6e-09;				
Matches 174; Conservative	0; Mismatches 96; Indels 19; Gaps 2;				
Oy	901	gttgggcaactcttcctgtgaagaacagacaggaactatttaagctcgtgtgcgat	960		
Db	103	GGTCAGCAAACTTTTGCTTAAGGGCATCAAAATCATTATTTCAGTCTGCGAGACAT	162		
Oy	961	atggtctcaagtacaactaccatcatctcgccctctgtagcacgaaagcaattagacaacat	1020		
Db	163	TTCGCTTTGCTGGAACACTCTCACTCTGCTACATAACAAAAGCCACAGACAT	222		
Oy	1021	atgcacaacatatgtgacccccctgaaaccttatttat---taggatacggaaacc	1077		
Db	223	ATGTAACAACAAAGAGGCTGACTGTAAATAAAGTATTATTATGACACACCTGAATTTGAAC	282		
Oy	1078	tgaaaataatgctttct-----tttgatttttcccacaataatta	1121		
Db	283	TCAAAACATTTTCATGCTTCCCACAAAACATTTCTACTTTTGACTATTTTCCAACCATTTAA	342		
Oy	1122	aacgtaaaaactactcttagtgcgaaggtaagccattcctaagctag	1170		
Db	343	AAAGTAACAACCATTTCTTAGGTGACAGCGCTGAGAAAAAACAATGTCATAG	391		
RESULT	9				
LOCUS	W76494 /c W76494 411 bp mRNA linear EST 17-OCT-1996				

```

DEFINITION      2467d03.r1 Soares_fetal_heart_NbHn19w Homo sapiens cDNA clone
ACCESSION       W76494
VERSION         W76494.1
KEYWORDS        GI:1386739
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE       1 (bases 1 to 411)
AUTHORS        Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,D.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterson
R., Williamson,A., Wobldmann,P. and Wilson,R.
The Washu-Merck EST Project
JOURNAL         Unpublished (1995)
COMMENT         Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wuston.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
Insert Length: 466 Std Error: 0.00
Seq primer: mob.REGA+ET.
FEATURES
SOURCE
1..411
/organism="Homo sapiens"
/db_xref="GDB:1271076"
/db_xref="taxon:9606"
/clone="IMAGE:345701"
/clone_lib="Soares_fetal_heart_NbHn19w"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT73 (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTCACATCTGAAATGGAGGAGCGGCCGACATCTTTTCTTTTCTTTT 3'), RI
double-stranded cDNA was size selected, ligated to Eco RI
adapters (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Felima Bonaldo. This library was constructed from the
same fecus as the fetal lung library, Soares fetal lung
NbHn19w."
BASE COUNT      129 a 58 c 79 g 145 t
ORIGIN
Query Match 3.4%; Score 77.4; DB 10; Length 411;
Best Local Simlarity 60.2%; Pred. No. 5, 6e-09;
Matches 174; Conservative 0; Mismatches 96; Indels 19; Gaps 2.
Db 901 gtggagcaactcttcctctaaagcaacagagacattatttagcctgtgcat 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
325 GGTACGCAAACTTTTCTGTGAAGGCCATACATATCATATTTTCACTCTGCAGACAT 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
961 atgctcagtcacactactactctctgcctctgtagcaggaattgcaacaat 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
265 TTGGCTTTTGGCTGAACCTACCTCACTGCTGTAACATACACAAANAACCTGCCAGACAT 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1021 atgtcaacaacatagtgaacccatggaacattttat---tagtatacggaaac 1077
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
205 ATGTAAACAAGAGGGGTACTGTATTAAGATGTATTTATGACACACTGAATTTTAAC 146
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1078 tgaataataatgctctct-----tttgattttttccccaatcatnaaa 1121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
145 TCAATATCATTTTGATGCTTCCCAAAACATTTCACTTTTGGACGATTTTTCACACATTAA 86

```



ACCESSION A1240516  
 VERSION A1240516.1 GI:3835913  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 322)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([infoimage.lnl.gov](http://infoimage.lnl.gov)) for further information.  
 Insert length: 735 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 315.  
 Location/Qualifiers  
 1..322  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1ib="IMAGE:1844499"  
 /lab\_host="DH10B"  
 /note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH119W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 90 a 72 c 67 g 93 t  
 ORIGIN  
 Query Match 3.4%; Score 77; DB 9; Length 322;  
 Best Local Similarity 71.6%; Pred. No. 6.6e-09;  
 Matches 101; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 Oy 891 ctgagccaggttggaacactctctcgtlaaagaacagaggaacttttagctc 950  
 |||||||  
 Db 178 ctgagccagaggttggaacactctctcgtlaaagaacagaggaacttttagctc 119  
 Oy 951 tctgtgcataatgctcagtcacaaactcactcctcctctgtagcagaagcaat 1010  
 |||||||  
 Db 118 tctgtgcataatgctcagtcacaaactcactcctcctctgtagcagaagcaat 59  
 Oy 1011 tagcaacaatattgtcaacaaa 1031  
 |||||||  
 Db 58 catgagatggttaagtcacacaaa 38  
 RESULT 13  
 AA443938 386 bp mRNA linear EST 03-JUN-1997  
 LOCUS zv51c03.s1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:757156  
 DEFINITION 3', mRNA sequence.  
 ACCESSION AA443938  
 VERSION AA443938.1 GI:2156613  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 386)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.  
 TITLE Washington University School of Medicine  
 JOURNAL 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 COMMENT Contact: Wilson RK  
 Unpublished (1997)  
 REFERENCE 1 (bases 1 to 321)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([infoimage.lnl.gov](http://infoimage.lnl.gov)) for further information.  
 Insert length: 735 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 321.  
 Location/Qualifiers  
 1..386  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1ib="IMAGE:757156"  
 /lab\_host="DH10B"  
 /note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer (5' TGTTACCAATCGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot1, and was constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 111 a 72 c 68 g 135 t  
 ORIGIN  
 Query Match 3.4%; Score 77; DB 9; Length 386;  
 Best Local Similarity 63.0%; Pred. No. 7e-09;  
 Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
 Oy 881 ttcccatgctcgtggaacaggttggaacactctcgtlaaagaacagaggaacta 940  
 |||||||  
 Db 139 ttgtttctctcgtggaacaggttggaacactctcgtlaaagaacagaggaacta 198  
 Oy 941 tttagagctctgtgtccatagtcacaaactcactcctcctctgtagca 1000  
 |||||||  
 Db 199 tttagagctctgtgtccatagtcacaaactcactcctcctctgtagca 258  
 Oy 1001 cgaagcaattagcaacaatattgtcaacaaactatgtgagcccatgaaacttatt 1060  
 |||||||  
 Db 259 cagaaactcgcacagacaaactatgtgagcccatgaaacttatt 318  
 Oy 1061 ttatgata 1069  
 |||||||  
 Db 319 attttccaaa 327  
 RESULT 14  
 AA444117/c 393 bp mRNA linear EST 03-JUN-1997  
 LOCUS zv51c03.r1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:757156  
 DEFINITION 5', mRNA sequence.  
 ACCESSION AA444117  
 VERSION AA444117.1 GI:2156792  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 393)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schlienberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
 T., Waterston, R., and Wilson, R.  
 WashU-Merck EST Project 1997  
 Unpublished (1997)  
 Contact: Wilson RK

FEATURES	Location/Qualifiers
SOURCE	1. .393

```

/organism="Homo sapiens"
/db_xref="GDB:5977767"
/db_xref="taxon:9606"
/clone_1ld="IMAGE:757156"
/clone_1ld="Soares_testis_NHT"
/sex="male"
/lab_host="DH10a"
/notes="Vector: pRT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc. and primed with a Not I - oligo(df) primer [5].
TGGTACCAATCTCAAGTGGAGGCGGCCCAATTTTATTTTATTTT 3'.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization to Cot1, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
[www.bio.llnl.gov/bdnp/image/image.html](http://www.bio.llnl.gov/bdnp/image/image.html)  
Insert length: 605 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 412.  
location/Qualifiers  
1..417

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1837458"
/clone_1ib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/obj="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGACAGGAGCGCGCCAAATTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cots, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```

```
Search completed: August 16, 2002, 03:39:47
Job time: 10781 sec
```

RESULT	15
LOCUS	AI208768
DEFINITION	AI208768 417 bp mRNA linear EST 29-NOV-1998
ACCESSION	9398c10.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1837458
KEYWORDS	AI208768
SOURCE	AI208768.1 GI:3770710
ORGANISM	EST. human. Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 417)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .

---

1  
2  
3

4  
5  
6

7



QY 1076 ccgagaaataaagtcctcttccttgatctttccccaatcattaaaaagcaaaacac 1135  
 Db 477 TTTTAAATTCATATATATCTTTTTTGATATTTTGAACCAATTTAAAAATAAACAACA 418  
 QY 1136 tcttaagtcgcaaggttaagccattctcagcttaagcaatgacagcctgattgactgt 1195  
 Db 417 TTTCTTCTCTACAGGGCTATGCCA-----TAAACAGTCAGCAAGCCAGATTTGGCTAT 366  
 QY 1196 gaactacagttggccaatccctgatt 1221  
 Db 365 AGGCCACAGATATGCCAAACCCCTTTT 340

```

RESULT 2
US-10-027-632-119946/c
; Sequence 119946, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Mary, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119946
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-119946

```

Query Match	3.6%	Score 82.8	DB 7	Length 1007
Best Local Similarity	58.0%	Pred. No. 3.3e-12		
Matches 189	Conservative	0	Mismatches 127	Indels 10
				Gaps 2
QY 896	ccacaggtctggcacaacctctccctgtaaaagacacaggaacatactttaagctctgct	955		
Db 655	ccagagggcaggttaccttttctgtaaagagcagagatgaatatttttagcctttgtgag	596		
QY 956	gacataatgctctagtcacaactactcaatcctcgcctctgtgacacgaagaactatgca	1015		
Db 595	gccatfaaggtctctctgtgcactattatcaactttgctgttcagcgaatagcagccatag	536		
QY 1016	acaataatgcacaacacatctgtagccccaatgaaacattatctttaatgtagcggaaa	1075		
Db 535	atfaatncttfaaatgaaatgacatctgcttccatataaact--ttattttatggacactgaaa	478		
QY 1076	cccgaaataaactgctctctctcttgaatcttctcccaatcaatataaaaaacgtataaactac	1135		
Db 477	tttttaatttcattattattcttttttttgattttttgaacacattttttaaataaagaaacca	418		
QY 1136	tcttaagtcgcgaaggttaagcatctctcaagctctagtcagtagagctctgattgctctgt	1195		
Db 417	ttctttgctcacaaggcctatggcca-----ttaacagctcagcaaacgacatttgctctcat	366		
QY 1196	gacctacagcttgcgcacaatccctgatt 1221			

Db 365 AGGCCACAGTATGCCAACCCCTGTT 340

```

RESULT 3
US-10-105-299-12107/c
: Sequence 12107, Application US/10105299
: GENERAL INFORMATION:
: APPLICANT: Rosen, et. al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS950
: CURRENT APPLICATION NUMBER: US/10/105,299
: NUMBER OF SEQ. ID NOS: 15197
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12107
: LENGTH: 37437
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-105-299-12107

```

[illegible]

```

RESULT      4
US-10-105-299-12108/c
; Sequence 12108, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: P5950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12108
; LENGTH: 37442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12108

```

	Query Match	3.6%	Score 81.4	DB 7	Length 37442
	Best Local Similarity	63.0%	Pred. No. 4.1e-11		
	Matches 143	Conservative 0	Mismatches 11	Indels 3	Gaps 1
QY	809	tctctctttcttcacatcgcacgagctgtgaggcagttaccctgtcccccaggtcaaa	868		
DB	32329	ttccacttcattctttccatcttcttctttccagttatatttccctcaaaaaaacacacattaaaa	32270		
QY	865	----ttctcaactccggtttcccaatgcctcggacccaggttggggaacactcttcctgtaaga	925		
DB	32269	attgttccttttgcaattggtcagattatctaacacagggggctacggcaaaactttttctgaaaaga	32210		



```

OY 926 acgcagaggaactatlttagcctcgtgtgctgacatgctcagtcacaaactcatic 965
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32209 GCCAGATGGGTAATATATTTTAGCGCTTGTGGCCAAACATTCCTGCTGGCAACGTATTCATT 32150

OY 986 tctgcctctgtacgcagaaagcaatttagcaacaatcgtcacaac 1032
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32149 TCTGCTGCTATAGTACAAAAGCAGCATATAACAATAAGTAAAGCAAC 32103

RESULT 5
US-09-629-469A-12215/c
: Sequence 12215, Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OHSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 08435/0123
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12215
LENGTH: 5294
TYPE: DNA
ORGANISM: Homo sapiens
US-09-629-469A-12215

Query Match 3.5%; Score 81; DB 5; Length 5294;
Best Local Similarity 74.5%; Pred. NO. 2.2e-11;
Matches 102; Conservative 0; Mismatches 35; Indels 0; Gaps 0

OY 894 gaccacagtggtggcaaacctctcgtttaaagacacagacgaactcttttagcctcgt 953
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1915 GACCAGAGATCAACAAAACCTTTCTGTAAAGGCCAGACAGTAAGTATTTTAAAGGTTTGT 1856

OY 954 gtgcacatcgtctcagtcacaaactacatcctcgtcctcgtgacacgaagaattag 1013
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1855 GCGGCATATGCTGCTGTGGCAATTACTTAATCTTGCCATCATATAGCAGAAAGCAGCTAC 1796

OY 1014 caacacatcgtcaaca 1030
      ||||| ||||| |||||
Db 1795 AGACCAATATGTACGCA 1779

RESULT 6
US-09-918-995-12893/c
: Sequence 12893, Application US/09918995
GENERAL INFORMATION:

```

```

1  APPLICANT: Hyseq, Inc.
2  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
3  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
4  FILE REFERENCE: 20411-756
5  CURRENT FILING DATE: US/09/918, 995
6  PRIOR APPLICATION NUMBER: 2001-07-30
7  PRIOR FILING DATE: 1999-01-20
8  NUMBER OF SEQ. ID NOS: 38054
9  SOFTWARE: FASTSEQ for Windows Version 3.0
10 SEQ ID NO 12893
11 LENGTH: 503
12 TYPE: DNA
13 ORGANISM: Homo sapiens
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: (1)..(503)
17 OTHER INFORMATION: n = A,T,C or G
18 US-09-918-995-12893
19
20 Query Match 3.5%; Score 79.4; DB 5; Length 503;
21 Best Local Similarity 58.4%; Pred.No.2,1e-11;
22 Matches 157; Conservative 0; Mismatches 111; Indels 1; Gaps 1
23
24 Qy 887 atgcctgagccacgagttggtggcaaaccttcctcgtlaagaacagacaggaacttlltag 946
25 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
26 Db 459 ATGGCTTCACATCCATCGGTCACGACGACACTCTCTGTAGTCGCCAGCTAGTACTAGTTAG 400
27
28 Qy 947 gctcttgtyccatatggtctcagtlcaacaactacatctctgcctctgtagcagaaag 1006
29 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
30 Db 399 GCTCTGTTGGGCGCACACATCTTGGTTACGCTACCTACGCTGCGCCCTGTAGCATGTACAG 340
31
32 Qy 1007 caattagcaacaatgatgtcaacaacatatgtgaccoccatlgaaaacttattattatgt 1066
33 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 Db 339 CAGTCACACAGATATATGTCAATGATGAATGATGCTGCGCCCATTTTAAACAATTTATA 280
35
36 Qy 1067 atacggaacccg-aaaataatgctcttccttlltgatttttcccaacalctlaaaaaag 1125
37 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
38 Db 279 ACAATATGACATATGTCAACAAATATATTTCTTTGGATTTTTTTTTCACGATTTTAAAGTA 220
39
40 Qy 1126 taataaactactcttagtgatgcgaagttaa 1154
41 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
42 Db 219 TAAATAACACAAACATGAAAAAATCTTAA 191
43
44 RESULT 7
45 US-09-785-276A-55453/C
46 Sequence 55453; Application US/09785276A
47 GENERAL INFORMATION:
48 APPLICANT: Schlegel, Robert
49 APPLICANT: Endege, Wilson
50 APPLICANT: Monahan, John
51 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
52 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
53 TITLE OF INVENTION: HUMAN PROSTATE CANCER
54 FILE REFERENCE: MRI-007B
55 CURRENT APPLICATION NUMBER: US/09/785, 276A
56 CURRENT FILING DATE: 2001-02-16
57 PRIOR APPLICATION NUMBER: 60/2183, 319
58 PRIOR FILING DATE: 2000-02-17
59 PRIOR APPLICATION NUMBER: 60/189, 862
60 PRIOR FILING DATE: 2000-03-16
61 PRIOR APPLICATION NUMBER: 60/207, 454
62 PRIOR FILING DATE: 2000-05-25
63 PRIOR APPLICATION NUMBER: 60/211, 314
64 PRIOR FILING DATE: 2000-06-09
65 PRIOR APPLICATION NUMBER: 60/219, 007
66 PRIOR FILING DATE: 2000-07-18
67 PRIOR APPLICATION NUMBER: 60/255, 281
68 PRIOR FILING DATE: 2000-12-13
69 NUMBER OF SEQ ID NOS: 62232
70 SOFTWARE: FastSeq for Windows Version 4.0

```

```
; SEQ ID NO 55453
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 561..582
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-55453
```

```
Query Match          3.5%: Score 79.4; DB 5; Length 627;
Best Local Similarity 67.7%: Pred. No. 2.3e-11;
Matches 128; Conservative 0; Mismatches 56; Indels 5; Gaps 1;
```

```
Oy 907 caaaccttcctgtaagaacacagaaactatttgctgcgtggtacatgctc 966
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 476 CGAACTTTTCTGTAAAGCCAGAGTAATTTTCAGACTTTGTAGCCAAACACTC 417
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 967 tcaatcacactacatctctgctctctgtagacagaagaatagacaacatgta 1026
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 TCTGTACAAATATTCATCTGCGCATTTGTCCAGAAAGCCAGCACAATACACA 357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1027 acaaacatgtgac-----cccatgaaaactttattctatctgatacgsaaacctgaa 1081
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 ACAAGTAGACCTGCGCTGTCTCCAAATAAATTATTACAAACTAGAGCTAAGCGCTGAT 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1082 aataatgctc 1090
      |||||
Db 296 GATCGAGTC 288
```

```
RESULT 8
US-10-027-632-151080
; Sequence 151080, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151080
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151080
```

```
Query Match          3.5%: Score 79; DB 7; Length 749;
Best Local Similarity 72.0%: Pred. No. 3.2e-11;
Matches 103; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

```
Oy 885 ccattgcttgccacccaggttggaacactctctctgtaagaagccagagaaactatttt 944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 ccgtgttagccagggctggggcaaaccttcacatgtaaaagccagatggttaataatttc 431
```

```
Oy 945 aggcctgtgcatatgctcagtcacaaactactctctgctctgtgtagcagaa 1004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 agactctgtggccatttgccctctgtcacaactactcagctctgtgctgtgtagcga 491
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1005 agcaattagcaacaatgtaaa 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 accagccacagacaataatgtaaa 514
```

```
RESULT 9
US-10-027-632-151081
; Sequence 151081, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151081
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151081
```

```
Query Match          3.5%: Score 79; DB 7; Length 749;
Best Local Similarity 72.0%: Pred. No. 3.2e-11;
Matches 103; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

```
Oy 885 ccattgcttgccacccaggttggaacactctctctgtaagaagccagagaaactatttt 944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 ccgtgttagccagggctggggcaaaccttcacatgtaaaagccagatggttaataatttc 431
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 945 aggcctgtgcatatgctcagtcacaaactactctctgctctgtgtagcagaa 1004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 agactctgtggccatttgccctctgtcacaactactcagctctgtgctgtgtagcga 491
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1005 agcaattagcaacaatgtaaa 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 accagccacagacaataatgtaaa 514
```

```
RESULT 10
US-10-027-632-7748/C
; Sequence 7748, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
```

```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7748
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-7748
```

```
Query Match 3.4%; Score 78.2; DB 7; Length 754;
Best Local Similarity 64.1%; Pred. No. 5,3e-11;
Matches 116; Conservative 1; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 892 tggaccaggttggcaaacctctctgttaagaacagaggaactatttagctct 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 239 TGGAGCAGGGCTTGGCAAACTTTCTGTAAAGGCAAGTGTAAATATTATTAGGCTTT 180

QY 952 tgtgacatattgtctcaagcaactcatctctgtcctctgtgacagaaagcaatt 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 GTGACCTTTCAGAGCTGTGTACAGACTCACTCACTGCTGCTGTGTGTAAGAAACAGCC 120

QY 1012 agcacacatatgtcaacaacatatgtgaccatgaaacttattattatgatalag 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 ACAGACAATAGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 60

QY 1072 g 1072
Db 59 g 59
```

```
RESULT 11
US-10-104-047-364
; Sequence 364, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 364
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-364
```

```
Query Match 3.4%; Score 77.2; DB 7; Length 2960;
Best Local Similarity 74.6%; Pred. No. 1,9e-10;
Matches 97; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 893 ggaccaggttggcaaacctctctgttaagaacagaggaactatttagctctg 952
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2470 ggaagccaggttgcacaacattctgttaagggccagatagtaaatatttggccttg 2529

QY 953 tttgcatattgtctcagtcacaactactcatctctgcctctgttagcagaagaacta 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2530 tgggcccattgtctctgtcacacagattcaactctgtgtgtaacaaaaagcagta 2589

QY 1013 gcaacaatat 1022
```

```
Db 2590 gacaatatgt 2599
```

```
RESULT 12
US-10-027-632-186700/c
; Sequence 186700, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186700
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-186700
```

```
Query Match 3.4%; Score 76.8; DB 7; Length 575;
Best Local Similarity 63.3%; Pred. No. 1,2e-10;
Matches 133; Conservative 1; Mismatches 73; Indels 3; Gaps 1;
```

```
QY 891 ctggaccaggttggcaaacctctctgttaagaacagaggaactatttagctc 950
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 CTGACCCAGGAGTAGGCAAACTTTCTTAAGGCCAGATAGTATATTATTTAGGATT 325

QY 951 tgtgcatattgtctcagtcacaactactcatctctgcctctgttagcagaagaact 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 TGGAGGCTTTCAGAGCTGTGTACAGACTCACTCACTGCTGCTGTGTGTAAGAAATCAGC 265

QY 1011 tagcaacaatatgtcaacaacatatgtgaccatgaaacttattattatgatalag 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CACAGATATATATATAAACAATGAGCATGCTGTGT---TCCAGTAACTATATGTCAC 208

QY 1071 ggaacctgaaatatatgtctcttttga 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 TGAATCTGAATTTAAACGCTTTCATGTCA 178

RESULT 13
US-10-027-632-142654/c
; Sequence 142654, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```

: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 142654
: LENGTH: 538
: TYPE: DNA
: ORGANISM: Human
US-10-027-632-142654
```

```
Query Match          3.4%; Score 76.6; DB 7; Length 538;
Best Local Similarity 64.7%; Pred. No. 1.3e-10;
Matches 130; Conservative 0; Mismatches 69; Indels 2; Gaps 1;
```

```
QY 901 gtgggcaaacctctcctgtaagaaccagacgaactattagctcgtgcat 960
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 214 GGTGGCAAACTTCTCTTAAGAGCAGTGAGTAACATTTTGGTTTCAGGCCAC 155
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 961 atggtctcagtcacacactcctcctcgtctgtagcagaagcaattagcaaat 1020
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 154 ATGCTCTGTGACCAACTACCTCCTGCACCTGAGCCTCAAGGCCACAGCAAT 95
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1021 atgtcaacaacatattgtagcccatgaaacttattt--atttgatacgaagaa 1078
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 94 ATTATAATTAGCAACGTGGCGCTCTTCATTAAACCTTCATTGTGTAACCTGAAATTT 35
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1079 gaaataatctctctcttctt 1099
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 34 GAAATTATATATATTTTCTATG 14
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 14
US-09-785-276A-55307/c
: Sequence 55307, Application US/09785276A
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Budege, Wilson
: APPLICANT: Monahan, John
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
: TITLE OF INVENTION: HUMAN PROSTATE CANCER
: FILE REFERENCE: MRI-0078
: CURRENT APPLICATION NUMBER: US/09/785,276A
: CURRENT FILING DATE: 2001-02-16
: PRIOR APPLICATION NUMBER: 60/183,319
: PRIOR FILING DATE: 2000-02-17
: PRIOR APPLICATION NUMBER: 60/189,862
: PRIOR FILING DATE: 2000-03-16
: PRIOR APPLICATION NUMBER: 60/207,454
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: 60/211,314
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/219,007
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/255,281
: PRIOR FILING DATE: 2000-12-13
: NUMBER OF SEQ ID NOS: 62232
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 55307
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-785-276A-55307
```

```

Query Match          3.4%; Score 76.6; DB 5; Length 608;
Best Local Similarity 74.0%; Pred. No. 1.3e-10;
Matches 97; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 901 gtgggcaaacctctcctgtaagaaccagacgaactattagctcgtgcat 960
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 183 GTTGGCAAACTTCTCTTAAGAGCAGTGAGTAACATTTTGGTTTCAGGCCAT 124
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 961 atggtctcagtcacacactcctcctcgtctgtagcagaagcaattagcaaat 1020
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 AAGCTCTTGTGACCAACTATTCATCTCTGCCATTTTGTGAGCAAGCCATAGCAAT 64
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1021 atgtcaacaacaa 1031
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 GATAAACAAA 53
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 15
US-10-027-632-233742/c
: Sequence 233742, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: FILE REFERENCE: 108927.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: PRIOR FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 233742
: LENGTH: 556
: TYPE: DNA
: ORGANISM: Human
US-10-027-632-233742
```

```

Query Match          3.3%; Score 75.6; DB 7; Length 556;
Best Local Similarity 65.9%; Pred. No. 2.4e-10;
Matches 108; Conservative 1; Mismatches 55; Indels 0; Gaps 0;

QY 901 gtgggcaaacctctcctgtaagaaccagacgaactatttagctcgtgcat 960
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 472 GGTGGCAAACTTCTCTTAAGAGCAGTGAGTAACATTTTGGTTTCAGGCCAT 413
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 961 atggtctcagtcacacactcctcctcgtctgtagcagaagcaattagcaaat 1020
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 412 ATGCTCTGTGACCAACTACCTGCTGACGACGTAAGCAAGCAATTTGTAACCAAT 353
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1021 atgtcaacaacatattgtagcccatgaaacttatttat 1064
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 352 ATGTAAGTAATGAAACATGGTGTGTTCAATAAATTTAATTT 309
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 16, 2002, 05:55:47  
Job time: 11641 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 02:31:36 ; Search time 485.09 Seconds  
(Without alignments)  
8091.007 Million cell updates/sec

Title: US-09-515-369b-1  
Perfect score: 2286  
Sequence: 1 taatacgaactacataaggg.....tgacttcacgcgcgtgagcgc 2286

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2286	100.0	2286	AAH26595	Human melanoma dif
2	81.8	3.6	37437	AAK84165	Human immune/haema
3	81.4	3.6	37442	AAK84166	Human immune/haema
4	81	3.5	5294	AAH14600	Human CDNA sequenc
5	77.6	3.4	26591	AAH36313	Human musculoskele
6	75.8	3.3	481	AAK74761	Human immune/haema
7	73	3.2	1837	ABA04430	Human PPI345 prote
8	71.6	3.1	628	AAH98619	Human EST-derived
9	71.4	3.1	286	AAH41353	Human secreted exp

C	10	71.2	3.1	407	21	AAC27788	Human secreted pro
C	11	70	3.1	123219	23	AAH88703	Human DNA sequence
C	12	69.8	3.1	50000	20	AAH23517	Human kidney amino
C	13	69.4	3.0	8029	22	AAH99747	Human protein enco
C	14	69.4	3.0	46765	24	AAH99306	DNA encoding Aldeh
C	15	69.4	3.0	47319	22	AAK64813	Human immune/haema
C	16	69.4	3.0	47319	22	AAK72230	Human immune/haema
C	17	69.4	3.0	236303	22	AAH11614	Human genomic DNA
C	18	68.6	3.0	325791	22	AAH3104	Human Oestrogen re
C	19	68.2	3.0	97662	22	AAH83908	Genomic sequence o
C	20	68	3.0	31140	22	ABA08065	Human ovarian and
C	21	68	3.0	31140	22	AAH06791	Human reproductive
C	22	67.8	3.0	17700	22	AAH68945	Human immune/haema
C	23	67.6	3.0	2515	22	AAH17444	Human granulocyte
C	24	67.6	3.0	6235	21	AAH29169	Human G-CSF genom
C	25	67.6	3.0	6679	21	AAH29170	Partial sequence o
C	26	67.6	2.9	6679	22	AAH17443	Human granulocyte
C	27	66.6	2.9	289	22	AAH39009	Novel human diago
C	28	66.4	2.9	466	22	AAH33971	Human colon cancer
C	29	66.4	2.9	655	22	AAH11797	Human breast cance
C	30	66	2.9	871	22	AAH04280	Human CDNA clone (
C	31	66	2.9	1712	22	AAH13903	Human CDNA sequenc
C	32	66	2.9	2291	22	AAH17522	Human CDNA sequenc
C	33	66	2.9	7496	22	AAK70490	Human immune/haema
C	34	66	2.9	7496	22	AAK74614	Human immune/haema
C	35	66	2.9	10225	22	AAH99021	Human excretory re
C	36	66	2.9	10225	22	AAH16371	Human kidney relat
C	37	66	2.9	24292	22	AAH15939	Human nervous syst
C	38	66	2.9	28731	22	AAK64887	Human immune/haema
C	39	66	2.9	28731	22	AAK69666	Human immune/haema
C	40	66	2.9	28731	22	AAK70491	Human immune/haema
C	41	66	2.9	28731	22	AAK74615	Human immune/haema
C	42	66	2.9	28731	22	AAK84942	Human immune/haema
C	43	65.6	2.9	872	22	AAH98871	Human EST-derived
C	44	65.4	2.9	519	22	AAH35469	Human musculoskele
C	45	64.8	2.8	15044	22	AAH36290	Human cardiovascular

## ALIGNMENTS

RESULT	1
ID	AAH26595 standard; DNA; 2286 BP.
XX	AAH26595;
AC	
XX	
DT	12-NOV-2001 (first entry)
DE	Human melanoma differentiation associated gene-7 (Mda-7) promoter.
XX	
KW	Melanoma differentiation associated gene-7; Mda-7; promoter; human
KW	neuroblastoma; astrocytoma; glioblastoma multiforme;
KW	cervical cancer; breast cancer; colon cancer; prostate cancer;
KW	osteosarcoma; chondrosarcoma; tumour; therapy; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	promoter
FT	Location/Qualifiers
FT	1..2240
FT	/*tag= a
FT	138..146
FT	/*tag= b
FT	/bound_motety= "C/EBP"
FT	446..454
FT	/*tag= c
FT	/bound_motety= "C/EBP"
FT	773..777
FT	/*tag= d
FT	/bound_motety= "C/EBP"
FT	801..807
FT	/*tag= e
FT	/bound_motety= "AP-1"

FT protein\_bind 1109..1113  
 FT /\*tag= f  
 FT /bound\_moiety= "C/EBP"  
 FT protein\_bind 1207..1214  
 FT /\*tag= g  
 FT /bound\_moiety= "AP1"  
 FT protein\_bind 1209..1214  
 FT /\*tag= h  
 FT /bound\_moiety= "C/EBP"  
 FT protein\_bind 1530..1537  
 FT /\*tag= i  
 FT /bound\_moiety= "AP1"  
 FT protein\_bind 2115..2121  
 FT /\*tag= j  
 FT /bound\_moiety= "AP1"  
 FT protein\_bind 2126..2133  
 FT /\*tag= k  
 FT /bound\_moiety= "C/EBP"  
 FT TATA\_signal 2210..2214  
 FT /\*tag= l  
 FT  
 FT MO200164921-A1.  
 PD 07-SEP-2001.  
 PD 28-FEB-2001: 2001MO-US06782.  
 PD 29-FEB-2000: 2000US-0515369.  
 PR (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA  
 PI Fisher PB, Madireddi MT:  
 PI WPI: 2001-565508/63.  
 PT  
 PT Melanoma differentiation associated gene-7 promoter capable of  
 PT treating cancer comprises directing transcription of heterologous  
 PT coding sequence encoding tumour suppressor polypeptide positioned  
 PT downstream, useful for treating cancer -  
 XX  
 XX Claim 1: Fig 6A-B; 132pp; English.  
 PS  
 XX  
 CC The present sequence is that of the promoter region of the human  
 CC melanoma differentiation associated gene-7 (Mda-7). It includes  
 CC 2240 bp from the 5' flanking region of the Mda-7 gene. The  
 CC promoter was isolated from a human placental genomic library  
 CC using a PCR-based method. The Mda-7 promoter exhibits melanocyte  
 CC tissue specificity and can only be activated in the targeted  
 CC tissue, i.e. the skin. Therefore, a gene of interest driven by  
 CC the Mda-7 promoter will be differentially expressed in these cells,  
 CC minimizing systemic toxicity. A recombinant expression construct  
 CC in which the human Mda-7 promoter is operably linked to a coding  
 CC sequence encoding a tumour suppressor protein is claimed. The  
 CC tumour suppressor is preferably p21, retinoblastoma protein or p53.  
 CC A host cell comprising the expression construct is also claimed.  
 CC And is preferably a tumour cell selected from a melanoma,  
 CC neuroblastoma, astrocytoma, glioblastoma multiforme, cervical  
 CC cancer, breast cancer, lung cancer or prostate cancer cell.  
 CC The promoter preferably comprises nucleotides 1-2240 of the present  
 CC sequence. A pharmaceutical composition including the recombinant  
 CC expression construct is used in a claimed method of treating  
 CC melanoma, neuroblastoma, astrocytoma, glioblastoma multiforme,  
 CC cervical cancer, breast cancer, colon cancer, prostate cancer,  
 CC osteosarcoma, chondrosarcoma or a cancer of the central nervous  
 CC system.  
 XX  
 XX Sequence 2286 BP; 547 A; 592 C; 510 G; 637 T; 0 other:

Query Match 100.0%; Score 2286; DB 22; Length 2286;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	taataagactactatagaagcgctgcagctcagatcacctcttgaaacccaggctcgtcgtcct	60
Db	1	taataagactactactaagaagcgctgcagctcagatcacctcttgaaacccaggctcgtcgtcct	60
Qy	61	ccaaagctctgactcaataactagatctccaaactgattgtggcgaagttcctaagttct	120
Db	61	ccaaagctctgactcaataactagatctccaaactgattgtggcgaagttcctaagttct	120
Qy	121	ctccttgacctcttcctcgaagaataatgatactgataagctcctcctcctcagaagaagctc	180
Db	121	ctccttgacctcttcctcgaagaataatgatactgataagctcctcctcctcagaagaagctc	180
Qy	181	aggcacatgttttcgtaagatccatccatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	240
Db	181	aggcacatgttttcgtaagatccatccatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	240
Qy	241	actcaagatcccaagtgatccctcgaagccagcaagtgatatacctgacctcgtcctc	300
Db	241	actcaagatcccaagtgatccctcgaagccagcaagtgatatacctgacctcgtcctc	300
Qy	301	gttcagagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	360
Db	301	gttcagagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	360
Qy	361	cagtagctcgaagatctcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	420
Db	361	cagtagctcgaagatctcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	420
Qy	421	gctgtcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	480
Db	421	gctgtcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	480
Qy	481	tcctcatcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	540
Db	481	tcctcatcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	540
Qy	541	gccctaaatccacatggtggaaggggaggtggtgggaaagaagtgctgcgtgtgggctgt	600
Db	541	gccctaaatccacatggtggaaggggaggtggtgggaaagaagtgctgcgtgtgggctgt	600
Qy	601	gccctaaatccacatggtggaaggggaggtggtgggaaagaagtgctgcgtgtgggctgt	660
Db	601	gccctaaatccacatggtggaaggggaggtggtgggaaagaagtgctgcgtgtgggctgt	660
Qy	661	gctatagccaaagcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	720
Db	661	gctatagccaaagcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	720
Qy	721	aataaacaacactagatgagatgagagggagggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	780
Db	721	aataaacaacactagatgagatgagagggagggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	780
Qy	781	gagctcgcgaaggggt	840
Db	781	gagctcgcgaaggggt	840
Qy	841	aggcagatcatcgttccccccagtcacatctctctctctctctctctctctctctctctctctct	900
Db	841	aggcagatcatcgttccccccagtcacatctctctctctctctctctctctctctctctctctct	900
Qy	901	gttggtcgaactctctcgttaagaagcaagcaagaaactatlttagctctgtgtgtgtgtgtgtgt	960
Db	901	gttggtcgaactctctcgttaagaagcaagcaagaaactatlttagctctgtgtgtgtgtgtgtgt	960
Qy	961	atgtgtcagtcagaactacatcatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1020
Db	961	atgtgtcagtcagaactacatcatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1020
Qy	1021	atgtcaaaaacataatgagcccatgaaactatattttatgtgatacgaagaactgtga	1080
Db	1021	atgtcaaaaacataatgagcccatgaaactatattttatgtgatacgaagaactgtga	1080
Qy	1081	aaataatgtcttctcttctgtatcttttccccaatcattaaaacacgttaaaaactactctta	1140



|||||  
Db 1081 aaaaagagctcttcttctgaatttttcccaatcaaaaagtaaaactacttca 1140  
QY 1141 ggtcgcaaggttaagcattctcagcttagcagtgagcagcttgatttgctgtgac 1200  
Db 1141 ggtcgcaaggttaagcattctcagcttagcagtgagcagcttgatttgctgtgac 1200  
QY 1201 acaagttagccaatccctgattcccaaaatglatctctcaaggatgtgagcaaatctat 1260  
Db 1201 acaagttagccaatccctgattcccaaaatglatctctcaaggatgtgagcaaatctat 1260  
QY 1261 gggagagtgctgagtaaacagagttaaagaagcaccagcaatcccaaggagctagcac 1320  
Db 1261 gggagagtgctgagtaaacagagttaaagaagcaccagcaatcccaaggagctagcac 1320  
QY 1321 atgcacaaggctctcaactcaactcaacttgattatctgtttataggagatcttgcaag 1380  
Db 1321 atgcacaaggctctcaactcaactcaacttgattatctgtttataggagatcttgcaag 1380  
QY 1381 acaagaattcccaaaaactagagcttgagagctgtgcttgaggaaaactgtctgtctg 1440  
Db 1381 acaagaattcccaaaaactagagcttgagagctgtgcttgaggaaaactgtctgtctg 1440  
QY 1441 atgcccctacatgggcacatgtagaatctagagctgagtgctgtgctgtagtgagatagg 1500  
Db 1441 atgcccctacatgggcacatgtagaatctagagctgagtgctgtgctgtagtgagatagg 1500  
QY 1501 gtcaagagctcttgacgtcccttgagcagcttgagacatcaagctgtctgtctccctgagt 1560  
Db 1501 gtcaagagctcttgacgtcccttgagcagcttgagacatcaagctgtctgtctccctgagt 1560  
QY 1561 ggttcagagccacaacagagccaaagactagcccaagagcaccagcctccagcttctg 1620  
Db 1561 ggttcagagccacaacagagccaaagactagcccaagagcaccagcctccagcttctg 1620  
QY 1621 ggcctgccaatggtatcttcttcttcttctgttccaaactaaaggagagcgac 1680  
Db 1621 ggcctgccaatggtatcttcttcttcttctgttccaaactaaaggagagcgac 1680  
QY 1681 atttggttgagtgatataaccctagagacatgggtagctgcatgctgagaaaacatc 1740  
Db 1681 atttggttgagtgatataaccctagagacatgggtagctgcatgctgagaaaacatc 1740  
QY 1741 ctcaactccttgccctgtatgagttaaaggaggaactacaaggttattcttcgctg 1800  
Db 1741 ctcaactccttgccctgtatgagttaaaggaggaactacaaggttattcttcgctg 1800  
QY 1801 ggaactatgccaagcagtagaaggagcattatgtaatttttgacaggaagagagac 1860  
Db 1801 ggaactatgccaagcagtagaaggagcattatgtaatttttgacaggaagagagac 1860  
QY 1861 tggctgtagctgcaagtaatagtatatacgtgtacagagtttacaagcacctct 1920  
Db 1861 tggctgtagctgcaagtaatagtatatacgtgtgtacagagtttacaagcacctct 1920  
QY 1921 gttgtgttttgccttgtttattacaacttgggacaatttttaaatatatacatgacg 1980  
Db 1921 gttgtgttttgccttgtttattacaacttgggacaatttttaaatatatacatgacg 1980  
QY 1981 agagctgtagcagagaagtaagtaagacttggccctggccacaagcagctgtgtagagcc 2040  
Db 1981 agagctgtagcagagaagtaagtaagacttggccctggccacaagcagctgtgtagagcc 2040  
QY 2041 tgaactcaaaacccaggtctatctcaactcagagggctgtcttcccatcgctgtattgac 2100  
Db 2041 tgaactcaaaacccaggtctatctcaactcagagggctgtcttcccatcgctgtattgac 2100  
QY 2101 cttaaaagttagtggtgactagggcaatgaaagtaattctctagaagaacatgaccaattcc 2160  
Db 2101 cttaaaagttagtggtgactagggcaatgaaagtaattctctagaagaacatgaccaattcc 2160  
QY 2161 ctcttccacctccctcttcttctccacccctcccccataagcccccataatagccc 2220  
|||||

Db 2161 ctcttccacctccctcttcttctccacccctcccccataagcccccataatagccc 2220  
QY 2221 aaatctccacaagccttgcttgccctgacaaaccttacttctgaatgatctccagcgt 2280  
Db 2221 aaatctccacaagccttgcttgccctgacaaaccttacttctgaatgatctccagcgt 2280  
QY 2281 gggagcg 2286  
Db 2281 gggagcg 2286  
RESULT 2  
AAK84165/c  
ID AAK84165 standard; DNA; 37437 BP.  
XX  
AC AAK84165;  
DT 07-NOV-2001 (first entry)  
XX  
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:38977.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
PN WO200157182-A2.  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
FE  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.



Db 32150 TTCTGCTAGTACAAAGCGCTATTAACAACTAGCAAC 32102  
RESULT 3  
AAK84166/C  
ID AAK84166 standard; DNA: 37442 BP.  
XX  
AC AAK84166;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38978.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001MO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 25-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 PR  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-483426/52.  
 DR  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Disclosure: SEQ ID NO 38978: 3071bp + Sequence Listing: English.  
 PS  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (1)  
 CC amino acid sequences given in AAM82170 to AAM91921. (1) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (1)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (1) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (1) by expressing inactive proteins or to  
 CC supplement the patient's own production of (1). Additionally, (1) `and  
 CC polynucleotides may be used to produce the secreted (1), by inserting the  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (1) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC  
 XX  
 XX Sequence 37442 BP; 10361 A; 6715 C; 6860 G; 13506 T; 0 other;  
 XX  
 Query Match 3.6%; Score 81.4; DB 22; Length 37442;  
 Best Local Similarity 63.0%; Pred. No.2.e-12;  
 Matches 143; Conservative 0; Mismatches 81; Indels 3; Gaps 1  
 QY 809 tctctcttcttcattcatcctccatcagctgagctgaggagcatcctctgccccagctcac 868  
 Db 32329 TCCACTTATCTCTTCAATGATGCTACTTCTTCCAGTATATATCTCTTAAACAACATPAAA 32270  
 QY 869 ---ttccctaccgccgttcccccagctcgagccaggtctggggaacctctctctgtaaga 925  
 Db 32269 ATTGTCCTTTGCAATGCTCAGTATATCTAAACAACAGGAGGCTCAGCAACTTTTCTGAAAATA 32210  
 QY 926 accagacagagaaactattttagctctctgtgtgcacatagtctcagtcacaaactacac 985  
 Db 32209 GCCAGATGTAATATTTTAAAGCTTTGTGGGCGACAGCATTCCTGTGTGCAACTATTTCATT 32150  
 QY 986 tctgctctctgtagcacgaagaagcaattagcaacaatactgtcaaaac 1032  
 Db 32149 TCTGCTCTATATAGTACAAAACGACCTATPAAACAAATTAAGTAAACCAAC 32103  
 RESULT 4  
 AAH14600/c  
 ID AAH14600 standard; cDNA: 5294 BP.

XX	AAH14600:
AC	
XX	
DT	26-JUN-2001 (first entry)
DE	Human CDNA sequence SEQ ID NO:12215.
XX	
KW	Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
OS	Homo sapiens.
PN	EP1074617-A2.
PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000EP-0116126.
XX	
PR	29-JUL-1999; 99JP-0248036.
PR	27-AUG-1999; 99JP-0300253.
PR	11-JAN-2000; 2000JP-0118776.
PR	02-MAY-2000; 2000JP-0183767.
PR	09-JUN-2000; 2000JP-0241899.
XX	
PA	(HELI-) HELIX RES INST.
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	Ishii S, Sugiyama T, Wakematsu A, Nagai K, Otsuki T;
DR	WPI: 2001-318749/34.
XX	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602
PT	full-length cDNAs defined in the specification, and for the detection
PT	and/or diagnosis of the abnormality of the proteins encoded by the
PT	full-length cDNAs -
XX	
PS	Claim 8; SEQ ID 12215; 2537pp + CD ROM; English.
XX	
CC	The present invention describes primer sets for synthesizing 5602
CC	full-length cDNAs defined in the specification. Where a primer set
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC	to the complementary strand of a polynucleotide which comprises one of
CC	the 5602 nucleotide sequences defined in the specification, where the
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	of an oligonucleotide comprising a sequence complementary to the
CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/3'-end sequence is selected from those defined in
CC	the specification. The primer sets can be used in antisense therapy and
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH3166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC	AAH95983 represent human amino acid sequences; and AAH13629 to AAH13632
CC	represent oligonucleotides, all of which are used in the exemplification
CC	of the present invention.
XX	
SQ	Sequence 5294 BP; 1637 A; 878 C; 1099 G; 1680 T; 0 other;
XX	
Query Match	3.5%; Score 81; DB 22; Length 5294;
Best Local Similarity	74.5%; Pred. NO. 9.5e-13;
Matches 102; Conservative	0; Mismatches 35; Indels 0; Gaps 0.
Oy	894 gaccacaggttggcacaactcttcgttaagaacacagcagaactattttagctctgt 953
Db	1915 GACGAGGAATCGAANAACCTTTTCTGCTAAAGGCCGACGACGTAAGTATTTTTACGCTTTCT 1856
Oy	gccccatagtctcagtcacactactcatcttcgtctgtatgcacgaaacgaattg 1013

Db 1055 GGCCATATGGTCTGTGGCAATPACTTACTTGCCATCATAGCAGAAAGCAGCTAC 1796  
Cy 1014 caacatactgcacaa 1030  
          |||||||  
Db 1795 AGACAATATGTACGCA 1779  
          "  
  
RESULT 5  
ID AAL36313/c  
ID AAL36313 standard; DNA; 26591 BP.  
XX  
AC AAL36313;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilicer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189674.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220563.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226581.  
PR 22-AUG-2000; 2000US-0226681.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228824.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.



PR	08-SEP-2000	2000US-023320801
PR	08-SEP-2000	2000US-02332081
PR	12-SEP-2000	2000US-02331968
PR	14-SEP-2000	2000US-02332997
PR	14-SEP-2000	2000US-02332998
PR	14-SEP-2000	2000US-02332999
PR	14-SEP-2000	2000US-02333000
PR	14-SEP-2000	2000US-02334011
PR	14-SEP-2000	2000US-02333063
PR	14-SEP-2000	2000US-02333064
PR	14-SEP-2000	2000US-02333065
PR	21-SEP-2000	2000US-02343723
PR	21-SEP-2000	2000US-02343724
PR	21-SEP-2000	2000US-02343977
PR	25-SEP-2000	2000US-02334998
PR	26-SEP-2000	2000US-02335498
PR	27-SEP-2000	2000US-02335834
PR	27-SEP-2000	2000US-02335836
PR	29-SEP-2000	2000US-02336127
PR	29-SEP-2000	2000US-02336127
PR	29-SEP-2000	2000US-02336368
PR	29-SEP-2000	2000US-02336370
PR	29-SEP-2000	2000US-02336370
PR	02-OCT-2000	2000US-02336802
PR	02-OCT-2000	2000US-02337037
PR	02-OCT-2000	2000US-02337038
PR	02-OCT-2000	2000US-02337039
PR	02-OCT-2000	2000US-02339340
PR	13-OCT-2000	2000US-02339345
PR	13-OCT-2000	2000US-02339337
PR	20-OCT-2000	2000US-02418221
PR	20-OCT-2000	2000US-02418221
PR	20-OCT-2000	2000US-02418260
PR	20-OCT-2000	2000US-02418187
PR	20-OCT-2000	2000US-02418187
PR	20-OCT-2000	2000US-0241809
PR	01-NOV-2000	2000US-02446117
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246569
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249219
PR	17-NOV-2000	2000US-0249220
PR	17-NOV-2000	2000US-0249221
PR	17-NOV-2000	2000US-0249222
PR	17-NOV-2000	2000US-0249223
PR	17-NOV-2000	2000US-0249224
PR	17-NOV-2000	2000US-0249225
PR	17-NOV-2000	2000US-0249226
PR	17-NOV-2000	2000US-0249227
PR	17-NOV-2000	2000US-0249228
PR	17-NOV-2000	2000US-0249229
PR	17-NOV-2000	2000US-0249230

PR	17-NOV-2000:	2000US-0249300.
PR	01-DEC-2000:	2000US-0250160.
PR	01-DEC-2000:	2000US-0250391.
PR	05-DEC-2000:	2000US-0251030.
PR	05-DEC-2000:	2000US-0251988.
PR	05-DEC-2000:	2000US-0256719.
PR	06-DEC-2000:	2000US-0251479.
PR	08-DEC-2000:	2000US-0251856.
PR	08-DEC-2000:	2000US-0251868.
PR	08-DEC-2000:	2000US-0251869.
PR	08-DEC-2000:	2000US-0251989.
PR	08-DEC-2000:	2000US-0251990.
PR	11-DEC-2000:	2000US-0254097.
PR	05-JAN-2001:	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI:	2001-483426/52.
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
XX		
PS	Disclosure; SEQ ID NO 29573;	3071bp + Sequence Listing; English.
XX		
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patient's own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK62169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
SQ	Sequence 481 BP: 140 A: 103 C: 79 G: 159 T: 0 other:	
	Query Match	3.3%; Score 75.8; DB 22; Length 481;
	Best Local Similarity	68.9%; Pred. No. 8.3e-12;
	Matches 104; Conservative	0; Mismatches 47; Indels 0; Gaps 0;
OY	912 tcttcctgtaagaaccagacaggaacatttttagcctctgtgagccatttgctcact	971
DB		
	282 ttttttctgtaaggccagacaggaataattttcaggcttggcgtgcgtgcttcgt	341
OY	972 cacacactactcaatctgctctgtagcaggaagaacatttagcacaacatactgcacaa	1031
DB		
	342 tgcacatgcctcaagttgcacatgttagaacacaaagcagcgtgatgatcaccacaa	401
OY	1032 catatgtgaccccatgaaacttatattatt	1062
DB		
	402 cgggtcgtgactgtgtcccaaaaaatttatt	432
RESULT	7	
ID	ABA04430	
AC	ABA04430: standard: cDNA: 1837 BP.	
XX		
DT	11-MAR-2002 (first entry)	
XX		





KW immunomodulatory; haematopoietic; chemokine; analgesic; hemostatic;  
 KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antihistaminic; vulnary; antiparkinsonian;  
 KW anticancer; osteopathic; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticoagulant; antidepressant; gene therapy;  
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200021990-A1.  
 XX  
 PD 20-APR-2000.  
 XX  
 PE 15-OCT-1999; 99WO-US24205.  
 XX  
 PR 15-OCT-1998; 98US-0104435.  
 XX  
 PA (GENET ) GENETICS INST INC.  
 PI Jacobs K, McCoy JM, LeVallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M;  
 DR MPI, 2000-317937/27.  
 XX  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (ESTs), useful for treating various disorders -  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 PS Claim 1; Page 198; 618bp; English.  
 XX  
 CC AAAA1261 to AAA4419 represent specifically claimed secreted expressed  
 CC sequence tags (ESTs), isolated from human, mouse, xenopus and rat  
 CC tissue sources. The ESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemoactive; analgesic; hemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antihistaminic; vulnary; anticancer; osteopathic; neuroprotective;  
 CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticoagulant; and antidepressant. The ESTs can be used for gene  
 CC therapy and in vaccines. The ESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the ESTs. Proteins encoded by the ESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumors, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA4320 to AAA4345 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 CC  
 CC  
 CC Sequence 286 BP; 88 A; 63 C; 50 G; 85 T; 0 other;  
 SQ  
 Query Match 3.1%; Score 71.4; DB 21; Length 286;  
 Best Local Similarity 72.1%; Pred.No.1.2e-10;  
 Matches 93; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
 QY 899 aggttggaacactcttcgtaagaacacagcaggacctttagctctgtygcc 958  
 Db 149 AGGCCAGATFAAACCTTTTCTGTAAAGACGACAACGTAATATTTAGCCCTGGTGTC 90  
 Y 959 atatgtcttcagtcaaacactactctctgcctctgtagcacgaagcaattgaca 1018

Db	89	ATATGCTCTCTGCGCACTACACAACTCTGCTTCTTGTATAAAGACGACATAGACA	30
Qy	1019	atatgtcaaa 1027	
Db	29	ATATGTAAA 21	
RESULT 10			
ID	AAc27788/JC		
XX	AAc27788 standard; cDNA; 407 BP.		
AC	AAc27788;		
XX			
DE	06-OCT-2000 (first entry)		
XX			
XX	Human secreted protein 5' EST, SEQ ID NO: 31863.		
KM	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;		
KW	gene therapy; chromosome mapping; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1033401-A2.		
PD	06-SEP-2000.		
XX			
PF	21-FEB-2000; 2000EP-0200610.		
XX			
PR	26-FEB-1999; 990S-0122487.		
XX			
PA	(GEST ) GENSET.		
PI	Dumas Mline Edwards J, Duclert A, Giordano J;		
XX	WPI: 2000-500381/45.		
DR			
XX			
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for		
XX	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for		
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures -		
PS	Claim 1; SEQ ID 31863; 71pp + CD-ROM; English.		
XX			
CC	The present sequence is one of a large number of 5' ESTs derived from		
CC	mRNAs encoding secreted proteins. No ORF has yet been conclusively		
CC	identified within the present sequence. The 5' ESTs were prepared from		
CC	total human RNAs or polyA+ RNAs derived from 30 different tissues. EST		
CC	sequences usually correspond mainly to the 3' untranslated region (UTR)		
CC	of the mRNA because they are often obtained from oligo-dT primed cDNA		
CC	libraries. Such ESTs are not well suited for isolating cDNA sequences		
CC	derived from the 5' ends of mRNAs and even in those cases where longer		
CC	cDNA sequences have been obtained, the full 5' UTR is rarely included.		
CC	5' ESTs are derived from mRNAs with intact 5' ends and can therefore be		
CC	used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used		
CC	in diagnostic, forensic, gene therapy and chromosome mapping procedures.		
CC	They are used to obtain upstream regulatory sequences and to design		
CC	expression and secretion vectors.		
XX			
SO	Sequence 407 BP; 118 A; 72 C; 84 G; 127 T; 6 other;		
Query Match 3.1%; Score 71.2; DB 21; Length 407;			
Best Local Similarity 63.6%; Pred. NO.1.7e-10;			
Matches 124; Conservative 0; Mismatches 69; Indels 2; Gaps 11.			
Qy	887	atgacctgaccacagattggacaactcttcgtataaagacacagacgaataattttag 946	
Db	402	ATGCCTAGACACACACACACGCGCAAAATTTTCTGTAAAGACGACAGATTAATATGTG 343	
Qy	947	gtctcgtgtgcataatggtctcaatgcaaacataactcatctctgcctctgtagacgaaag 1006	
Db	342	GCTTTGACAGGGCTACATGATCTCTGTATGACCAACACACACACACATGACCCCTCAATGTGAAAAA 283	
Qy	1007	caatgaacaacatattgtcaacaacacatattgacaccccatgaaacattatattatattat 1056	

Db 282 CAGCATTGAGATGCTCAACCAATGAGTGTGCTGTTCNCA--TACATTATTATGC 225  
1067 atacggaacactgaa 1081  
1 1111 1111  
Db 224 ACATTGAAATTTGAA 210

## RESULT 11

AAH8703 standard; DNA; 123219 BP.

AAH8703:

26-FEB-2002 (first entry)

Human DNA sequence SEQ ID 543.

Single nucleotide polymorphism; SNP; diallelic marker; human;  
central nervous system disorder; CNS; ds.

Homo sapiens.

MO200151659-A2.

19-JUL-2001.

11-JAN-2001; 2001WO-1B00116.

13-JAN-2000; 2000US-0175854.

(GEST ) GENSET.

Chu T, Blumenfeld M, Cohen D;

WPI; 2001-483085/52.

Isolated polynucleotides, useful for genotyping nucleic acids for  
diallelic markers for the diagnosis of depression, comprises central  
nervous system disorder related diallelic marker -

Disclosure: Page 439-472; 519pp; English.

The present invention relates to diallelic markers derived from human  
CNS involved in central nervous system (CNS) disorders (see  
AAH8161-AAH8702). The markers have a single nucleotide polymorphism  
(SNP) and are useful in determining the genetic predisposition of  
individuals to CNS disorders, by identifying the nucleotides at a set  
of genetic markers in a biological sample, where the markers comprise at  
least one CNS disorder related marker. The present sequence was used  
to illustrate the invention.

Sequence 123219 BP; 27921 A; 32327 C; 32369 G; 30437 T; 165 other:

Query Match 3.1%; Score 70; DB 23; Length 123219;

Best Local Similarity 61.5%; Pred. No. 1e-08;  
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

895 acccaggtggtggaacactcttcctgtaaaacagacaggaacatttgagctctg 954

17375 agcagggggtcagcaaacctttctgaaaggatagatgacactgcatagcttg 17434

955 tggcctatgtctcagtcacactctcctctgctcctctgagcagaaagcattagc 1014

17435 ggcacagtggtccatccatccatctcactctgcttgaagagaagcagcaca 17494

1015 aacaaatgctcaacaacatctgtgaccacgaaanaactttatattgatacgaa 1074

17495 gacaaatgctcacaagaagcggtgtctgcatgccaataaaacttattataaacat 17554

1075 ac 1076

Db 17555 ac 17556

## RESULT 12

AAAX23517/C  
ID AAAX23517 standard; DNA; 50000 BP.

AAAX23517;

23-JUN-1999 (first entry)

Human kidney aminopeptidase P genomic DNA fragment 1.

Aminopeptidase; human; Amp; gene therapy; treatment; Amp-deficiency;  
prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;  
arterial stenosis; industrial protein feed; malabsorption syndrome;  
proteinaceous waste degradation; additive; immunohistochemistry; ss.

Homo sapiens.

WO9911799-A2.

11-MAR-1999.

02-SEP-1998; 98WO-US18426.

02-SEP-1997; 97US-0057854.

(MEDT-) MEDICAL COLLEGE GEORGIA RES INST.

Ryan JW, Sprinkle JJC, Venema RC;

WPI; 1999-205193/17.

Nucleic acid encoding human aminopeptidase P

Claim 13; Page 80-109; 201pp; English.

This invention describes the isolation of a novel human aminopeptidase P  
(Amp). This protein is used to produce recombinant Amp and can be used  
for gene therapy for treating Amp-deficiency conditions. Its fragments  
are used as primers and probes to identify patients with homozygous and  
heterozygous Amp deficiency, including prenatal diagnosis (patients  
defective in Amp are at risk of developing angioedema if treated with  
angiotensin-converting enzyme inhibitors), also as antisense inhibitors  
in cases of excessive Amp expression. The product of the invention is  
also used to identify Amp-expressing sequences in other animals and to  
generate transgenic animals, and comparisons of genomic sequences are  
used to detect mutations. Amp inhibitors are potentially useful as  
antihypertensive agents and to prevent or treat arterial (re)stenosis  
or atherosclerosis. The structure of Amp is used to design synthetic  
substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal  
imido bonds, can be used to degrade industrial protein feeds to free  
amino acids, to degrade proteinaceous wastes, as additives in enzyme  
formulations used to treat malabsorption syndrome and for studying its  
biological role. Antibodies against Amp are used in immunohistochemical  
methods to study Amp distribution.

Sequence 50000 BP; 13187 A; 12125 C; 11842 G; 12846 T; 0 other:

Query Match 3.1%; Score 69.8; DB 20; Length 50000;

Best Local Similarity 72.7%; Pred. No. 7e-09;  
Matches 104; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

889 gctcggaccaggtgtgggcaacactcttcctgtaaaacagacaggaacattttgagc 948

8967 GTCAGATCAAGATTGGCAAA--CTTTCTTAAGGCTGAGATGATGATTTTGAC 8910

949 tctgtgtcatalatggtctcagtcacaactactatctgtcctctgtagacgaagca 1008

8909 TCTGCAGACATACAGTCTGTGTAGCAACCACTCACTCTGCTGTGCAGCCCAAGCA 8850

```
QY 1009 attagcaacatatgtcaaca 1031
      ||| ||||| | |||||
Db 8849 AGTATAGACAATATATAACAA 8827
```

RESULT 13

AAH99747	ID	AAH99747	standard; cDNA; 8029 BP.
XX	AC	AAH99747:	
XX	DF	16-OCT-2001	(first entry)
XX	DE	Human protein encoding cDNA sequence SEQ ID NO:582.	

KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KM antineoplastic; antineoplastic; antineoplastic; immunosuppressive;  
KM antibacterial; endocrine; cardiac; central nervous system; virucide;  
KM anti-HIV; fungicide; antitubercular; cardiovascular; antineoplastic; anaemia;  
KM antileprosy; haemostatic; vulnery; antileprosy; osteopathic; eczema;  
KM dermatological; antiallergic; antineoplastic; antidiabetic; cytosolic;  
KM neuroprotective; antidepressant; nootropic; antiparasitic; infection;  
KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KM antineoplastic; rheumatoid arthritis; septic shock; pancreatitis;  
KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KM genetic disease; haemopoietic disorder; platelet disorder; asthma;  
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KM allergic rhinitis; diabetes; multiple sclerosis; depression;  
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KM neurological disorder; ss-

OS Homo sapiens.

PN W0200153455-A2.

PD 26-JUL-2001

PF 22-DEC-2000; 2000WO-US35017.  
XY

PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000: 2000US-0488725

PR 25-APR-2000; 200005-0552317.  
XX  
XX

PA (HYSE-) HYSEQ INC.  
XX  
XX

XX Yang, YI, Liu C,  
XX

DR WP1; 2001-43/603/49.  
DR P-PSDB; AAM25806.

aa	Isolated human polypeptide
PT	

XX

[illegible]

CC AAM25963. The proteins can have activity

CC cells, they are expressed on the surface of haematopoietic, endothelial, epithelial, and certain tumour cells; CC anti-infective; immunosuppressive; antibacterial; endocrine; cardiac; CC central nervous system; vinorelbine; anti-HIV; fungicide; antimutagen; CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnereary; CC antiulcer; osteoplastic; dermatological; antiallergic; antiasthmatic; CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
SQ Sequence 8029 BP; 2064 A; 1942 C; 1905 G; 2118 T; 0 other;

Sequence 8029 BP; 2064 A; 1942 C; 1905 G; 2118 T; 0 other;

Query Match	3.0%	Score 69.4	DB 22	Length 8029
Best Local Similarity	57.2%	Pred. No. 3.2e-09		
Matches 202; Conservative	0	Mismatches 121	Indels 30	Gaps 3

QY 889 gcttggaccacaggtcttgggcaaacctcttcgcgttaaaagaccaggaacttctttagc 948  
Db 682 gctatgacacagagctggcaaaccttctctttaagagacagtagtaactcttttagc 741  
QY 949 tctgtgtgcataatgcttctcagttcaaacactcaactctctgtcctctgttagaac----- 1007  
Db 742 tttagcagcttactactgtctccctcccaacttcaactctctgcgaagaagaagctccacga 801  
QY 1002 -----gaaagcaacttagcaacaatatgttcaacaacaacataatgtgacccatga 1048  
Db 802 caataccaataatgatgtggtgtgtatgctcttaataaacattgttlttagaacatga 861  
QY 1049 aaacttattt-altataggatagagaaacctgaaataatgctcttctcttatttttc 1107  
Db 862 aatcctgaattctcaatgaattcttcaggtgtcaatgaataatgcttctctcttcttattctcg 921  
QY 1108 cccaactctaaanaagctaaanaactctcttggctgcacaggttaagcatcttcagct 1167  
Db 922 cccaatctattgnaaacaataataaacactcttagctcacag-----tltgataaa 972  
QY 1168 tagcaggtgcaggtctgtgatttggctgttgactcaagatagtcgcaatcccttcgat 1220  
Db 973 aacaggtgcaggtccaggttctgattctgtgtgtggccatagtgttgcacacccttcga 1025

RESULT 14

ID	standard; DNA; 46765 BP
xy	

AC AAS99306;  
YV

DT 12-MAR-2002 (first entry)  
XX

DE DNA encoding Aldenhyde dehydrogenase 5 family, member A1 (ALDH5A1)

KW succinate-semialdehyde dehydrogenase: gene therapy;  
KW Aldehyde dehydrogenase 5 family member A1; ALDH5A1;

4-hydroxybutyric aciduria; metabolic disease; transgenic animal; antisense technology; allele specific oligonucleotide; ASO; KM

cytosome opz; us.  
xx  
xx  
xx

XX  
XX  
XX

FT variation

PT

ET

FT variation

**EJ**

ET  
ET

FT	variation
FT	

variation  
FJ  
FJ

333

location/Qualifiers	
replace(3749,T)	
*tag= a	"Single nucleotide polymorphism"
/standard_name=	
replace(3800,G)	
*tag= b	"Single nucleotide polymorphism"
/standard_name=	
replace(3829,T)	
*tag= c	"Single nucleotide polymorphism"
/standard_name=	
replace(3875,G)	
*tag= d	"Single nucleotide polymorphism"
/standard_name=	
replace(3876,A)	
*tag= e	"Single nucleotide polymorphism"
/standard_name=	
replace(3941,T)	
*tag= f	"Single nucleotide polymorphism"
/standard_name=	
	"Single nucleotide polymorphism"

```

FT      exon      4050..4403
FT      /tag- g
FT      /number- 1
FT      4050..42765
FT      /tag- h
FT      /product- "ALDH5A1"
FT      /note- "Aldehyde dehydrogenase 5 family member A1.
FT      Specifically claimed in claim 25"
FT      replace(4059,G)
FT      /tag- i
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(4111,A)
FT      /tag- j
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(4155,C)
FT      /tag- k
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(4179,G)
FT      /tag- l
FT      /standard_name- "Single nucleotide polymorphism"
FT      4404..11575
FT      /tag- m
FT      /number- 1
FT      replace(4430,G)
FT      /tag- n
FT      /standard_name- "Single nucleotide polymorphism"
FT      11576..11659
FT      /tag- o
FT      /number- 2
FT      11660..12315
FT      /tag- p
FT      /number- 2
FT      12316..12486
FT      /tag- q
FT      /number- 3
FT      replace(12415,T)
FT      /tag- r
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(12422,T)
FT      /tag- s
FT      /standard_name- "Single nucleotide polymorphism"
FT      12487..13921
FT      /tag- t
FT      /number- 3
FT      replace(12559,T)
FT      /tag- u
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(13855,G)
FT      /tag- v
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(13917,T)
FT      /tag- w
FT      /standard_name- "Single nucleotide polymorphism"
FT      13922..14038
FT      /tag- x
FT      /number- 4
FT      replace(14021,T)
FT      /tag- y
FT      /standard_name- "Single nucleotide polymorphism"
FT      14039..24219
FT      /tag- z
FT      /number- 4
FT      replace(14046,A)
FT      /tag- aa
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(14179,A)
FT      /tag- ab
FT      /standard_name- "Single nucleotide polymorphism"
FT      24220..24363
FT      /tag- ac
FT      /number- 5
FT      24364..29453
FT      /tag- ad

```

```

FT      /number- 5
FT      24598..31819
FT      /tag- ae
FT      /number- 6
FT      replace(29397,T)
FT      /tag- af
FT      /standard_name- "Single nucleotide polymorphism"
FT      29454..29597
FT      /tag- ag
FT      /number- 6
FT      replace(29469,A)
FT      /tag- ah
FT      /standard_name- "Single nucleotide polymorphism"
FT      31820..31978
FT      /tag- ai
FT      /number- 7
FT      31979..37049
FT      /tag- aj
FT      /number- 7
FT      37050..37219
FT      /tag- ak
FT      /number- 8
FT      37220..41171
FT      /tag- al
FT      /number- 8
FT      41172..41230
FT      /tag- am
FT      /number- 9
FT      replace(41217,C)
FT      /tag- an
FT      /standard_name- "Single nucleotide polymorphism"
FT      41231..42559
FT      /tag- ao
FT      /number- 9
FT      replace(41237,A)
FT      /tag- ap
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(41403,C)
FT      /tag- aq
FT      /standard_name- "Single nucleotide polymorphism"
FT      replace(42380,T)
FT      /tag- ar
FT      /standard_name- "Single nucleotide polymorphism"
FT      42560..42765
FT      /tag- as
FT      /number- 10
FT      exon
FT      WO200190119-A2.
FT      29-NOV-2001.
FT      21-MAY-2001; 2001WO-US16558.
FT      19-MAY-2000; 2000US-205849P.
FT      (GENA-) GENAISSANCE PHARM INC.
FT      Kilem SE, koshy B, Tanguay DA.
FT      WPI: 2002-089912/12.
FT      P-PSDB: AA073594.
FT      DR
FT      New genetic variants of human aldehyde dehydrogenase 5 family, member
FT      ALDH5A1 gene for treating metabolic diseases and for expressing
FT      ALDH5A1 protein useful in identifying drugs to treat 4-hydroxybutyric
FT      aciduria
FT      Disclosure: Fig 1; 151pp; English.
FT      The invention describes an isolated polynucleotide comprising a
FT      nucleotide sequence which is a polymorphic variant of a reference
FT      sequence for the aldehyde dehydrogenase 5 family, member A1
FT      (succinate-semialdehyde dehydrogenase) (ALDH5A1) gene or its fragment.

```

CC The polypeptide is useful for screening for drugs targeting it by  
CC contacting the ALDH5A1 polymorphic variant with a candidate agent and  
CC assaying for binding activity. The polypeptide and haplotypes are useful  
CC for identifying an association between a trait such as a clinical  
CC response to a drug targeting ALDH5A1 and a haplotype ALDH5A1 gene.  
CC Transgenic animals are also useful for studying expression of the ALDH5A1  
CC isogenes in vivo, for in vivo screening and testing of drugs against  
CC ALDH5A1 protein and for testing the efficacy of therapeutic agents and  
CC compounds for 4-hydroxybutyric aciduria and metabolic diseases in a  
CC biological system. Antibodies are useful for diagnostic and prognostic  
CC formats and therapeutic methods, for immunoprecipitating the polypeptide  
CC from solution, for detecting ALDH5A1 protein isoforms in biological  
CC samples, frozen tissue sections, for use in immunocytochemical,  
CC immunohistochemical and immunofluorescence techniques. The polynucleotide  
CC is useful for gene therapy and antisense gene therapy. This sequence  
CC encodes ALDH5A1 (located on chromosome 6p22), and forms a reference  
CC sequence on which polymorphic sequences encoding isoforms of the ALDH5A1  
CC protein are based, described in the method of the invention.

Query Match 3.08; Score 69.4; DB 24; Length 46765;  
Best Local Similarity 57.28; Pred. No. 8.8e-09;  
Matches 202; Conservative 0; Mismatches 121; Indels 30; Gaps 3;

OY 889 gccggaccaggtgtggcaaaccttcctgtaagaagaccagagcaactatttaggc 948  
DB 45589 GCTTAGATCAGGAGCTGGCAAACTTTTCTTAAGAGACAACTAGTAATCTTTAGGC 45530  
OY 949 tctgtgtgcatacgtctcagtcacaactaactcctcgtcctcgttagac----- 1001  
DB 45529 TTTGAGGCTACTGCTCCCTGCCAATTCACCTGCGCAAGAACAGCTCCAGA 45470  
OY 1002 -----gaaagcaattggcaacaaitgtcaacaacaacatcgtgaccacatga 1048  
DB 45469 CAATACCAATTAATGATGGGTGGCTATGCTTAATTAACATTTTGTATTGACACTGA 45410  
OY 1049 aaacttatct-attatgatacggaaacctgnaaataatgctctcttcttgaatttttc 1107  
DB 45409 AATCGAATTCATAGATTTTTCAGGTGTCATGAAATATTGTTCTTTGATTTCG 45350  
OY 1108 cccaacttaaaaaagtaaaactactccttagtctgcgaagttaagccatctcaagc 1167  
DB 45349 CCCAATCATTTGAAACATATTAACCAATCTTACCTCAGG-----TTGTATMAA 45299  
OY 1168 tagcagtgagcagctgattgtgcttgcactacagttgccaactcctgat 1220  
DB 45299 AACAGGTGGCAGGCCAGGCTTTGATTCGTGGGCCATAGTTTGCACCCCTGAT 45246

RESULT 15  
AAK64813/C  
ID AAK64813 standard; DNA; 47319 BP.  
XX  
AC AAK64813;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19625.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
XX  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX  
XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216687.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 22-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.

02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0244675.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251890.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 19625; 3071pp + Sequence Listing; English.  
XX

CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins, and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
-SQ Sequence 47319 BP; 13708 A; 8453 C; 9720 G; 15438 T; 0 other;

Query Match 3.08; Score 69.4; DB 22; Length 47319;  
Best Local Similarity 66.28; Pred. No. 8.9e-09;  
Matches 100; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 881 ttcccatgctgacccaggttggaactcttcctgtaagaacagaggaacta 940  
DB 18549 TTTACTAAGTCTAAATCAGACATGGCAACTTTTCTGTCAGACGCCAGCATTAATTA 18490  
OY 941 tttagctctgctgcatatgltcagtcacaactactactactctgctctgtaga 1000  
DB 18489 TTTAGGCTTTGTAAAGCCACAGCTCTCTACGGCACACCACTCAACTCTGCATATAGCC 18430  
OY 1001 ggaagcaaltagcacacatatgtccaacaa 1031  
DB 18429 CCAAGCTCTACAGATGATTAATAATGAA 18399

Search completed: August 16, 2002, 05:33:21  
Job time: 10905 sec

Mon Aug 19 06:46:33 2002

us-09-515-369b-1.rng

Page 17





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 03:40:03 : Search time 5721.79 Seconds

(without alignments)  
8643.690 Million cell updates/sec

Title: US-09-515-369B-1

Perfect score: 2286  
1 taatagcactacatagggg.....tgacttcacgctcggagc 2286

Scoring table: IDENTITY,NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending-Patents-NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/US006\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq:\*  
17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:\*  
20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq:\*  
21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq:\*  
22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq:\*  
23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq:\*  
24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq:\*  
25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:\*  
26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:\*  
27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:\*  
28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:\*  
29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq:\*  
30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq:\*  
31: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*  
32: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*  
33: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*  
34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*  
35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*  
36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*  
37: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq:\*  
38: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq:\*  
39: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq:\*  
40: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq:\*  
41: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq:\*  
42: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq:\*  
43: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq:\*

SUMMARIES					Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
Result No.	Score	Query Match	Length DB	ID	Description	
1	2286	100.0	2286	19	US-09-515-369B-1	
2	420	18.4	426	19	US-09-528-409-108247	
3	420	18.4	426	35	US-09-933-524-108247	
4	420	18.4	426	35	US-09-933-524A-108247	
5	93	4.1	338	1	PCT-US01-15674A-421	
6	93	4.1	338	22	US-09-573-080A-421	
7	87	3.8	196093	63	US-60-245-228-200	
8	87	3.8	196093	64	US-60-258-272-80	
9	86	3.8	224	1	PCT-US01-15674A-420	
10	86	3.8	224	22	US-09-573-080A-420	
11	86	3.8	76534	63	US-60-243-468-358	
12	83.2	3.6	179040	61	US-60-226-176-1401	
13	83.2	3.6	179040	62	US-60-233-468-1401	
14	83.2	3.6	179040	70	US-60-313-371-1401	
15	83	3.6	163174	60	US-60-212-664-256	
16	82.8	3.6	1007	24	US-09-634-3068-119945	
17	82.8	3.6	1007	24	US-09-634-3068-119946	
18	82.4	3.6	17947	60	US-60-070-694-1633	
19	81.8	3.6	288	14	US-09-044-030-1633	
20	81.8	3.6	288	43	US-60-172-360-1441	
21	81.8	3.6	511	56	US-60-185-216-346	
22	81.8	3.6	511	57	US-60-185-216-346	
23	81.8	3.6	37437	1	PCT-US01-01354-38977	
24	81.8	3.6	37437	30	US-09-764-905-38977	
25	81.8	3.6	37437	35	US-09-950-083-9956	
26	81.8	3.6	37437	37	US-10-092-399-38977	
27	81.4	3.6	37442	1	PCT-US01-01354-38978	
28	81.4	3.6	37442	30	US-09-764-905-38978	
29	81.4	3.6	37442	35	US-09-950-083-9957	
30	81.4	3.6	37442	37	US-10-092-399-38978	
31	81.2	3.6	161395	61	US-60-226-176-1603	

## ALIGNMENTS

ORGANISM: Human  
US-09-515-369B-1

Best Local Similarity 100.0%; Pred. No. 0  
Matches 2286; Conservative 0; Mismatches

oy 421 gggtctctctgctcccaaccataatactcagcaatctcaactgtctctcttctg 480  
|||||

Db 1501 gtcaagagctcttgactgcgccttgcaagctcttgaacacatcaagctgtctcgtgtcccccagagt 1560



```
Db 365 GMAAAGCTGGGCTCCAGAGAAAGGATTCAGGCTGGGCGAGCTATAGCCAGACAG 306
Oy 676 actgttgccaggagatgcaaaaggatatttctgttcttaagaaataaacaactcga 735
Db 305 ACTGCTGCCAGGATTTCCAAAGAGATTTTGTGTTGAAGAAATAAACACACGCA 246
Oy 736 gtaagatggaaggagggtgtgtgtgcaagagagatctggaaagatctgccaaggt 795
Db 245 GATAGATGAGAGGAGGAGGCTGTGTGCGAGAGAGATGGGAAGACTGCCAAGGCT 186
Oy 796 gttcttctactctctctcttcttcttcttcttcttcttcttcttcttcttctt 855
Db 185 GGTGTTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
Oy 856 ccccaagctcaatcttcttcttcttcttcttcttcttcttcttcttcttcttct 915
Db 125 CCCCACGTCACATTCCTACTCCGTTTCCATCCCTGGACCCAGCTTGGCCAACTCT 66
Oy 916 cctgttaagaaccagagaaactatatttagctctgtgtgcatatggtctcaatc 975
Db 65 CCTGTAAGAACACAGAGAACTATTGAGGCTGTGTGTCATATGCTCAGTCACA 6
```

```
RESULT 4
US-09-933-524A-108247/C
: Sequence 108247, Application US/09933524A
: GENERAL INFORMATION:
: APPLICANT: Drmanec, Radolje T.
: APPLICANT: Labat, Ivan
: APPLICANT: Stache-Crain, Birgit
: APPLICANT: Dickson, Mark
: APPLICANT: Jones, Lee W.
: TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
: TITLE OF INVENTION: From Various Libraries
: FILE REFERENCE: 774
: CURRENT APPLICATION NUMBER: US/09/933,524A
: CURRENT FILING DATE: 2001-08-20
: PRIOR APPLICATION NUMBER: 09/526,409
: PRIOR FILING DATE: 2000-05-17
: NUMBER OF SEQ ID NOS: 116231
: SOFTWARE: Hy-Patent.pl Version 3.1
: SEQ ID NO 108247
: LENGTH: 426
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-933-524A-108247
```

Query Match 18.4%; Score 420; DB 35; Length 426;  
Best Local Similarity 100.0%; Pred. No. 2,1e-108;  
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 556 ggtggaaggggggagatggggaagagatgctgttgaggcttcttcttcttcttct 615
Db 425 GGTGGAGAGGGGAGTGGGGAGAGAGAGTGGCTGTGGGCTGCTCACTTCTGGAGG 366
Oy 616 gtaagatcgagccctccagaaagagatcagagctgtgtgcaagctataagccaa 675
Db 365 GTAAAGACTCGGCCCTCCAGAACAAAGAGATTCAGGCTGGTGCACCTATAGCAAG 306
Oy 676 actgttgccaggagatgcaaaaggatatttctgttcttaagaaataaacaactcga 735
Db 305 ACTGCTGCCAGGATTCGCAAGAGATTTTGTGCTTAAAGAAATAAACACACTGA 246
Oy 736 gtaagatggaaggagggtgtgtgtgcaagagagatgggaaagagcttgcgaaggt 795
Db 245 GTATGAGATGAGAGGAGGAGGCTGTGTGCGAGAGATTGGGAAGAGTCTGCCAAGG 186
Oy 796 gttcttactactcttcttcttcttcttcttcttcttcttcttcttcttcttctt 855
Db 185 GTGTTCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
```

```
Oy 856 ccccaagctcaatcttcttcttcttcttcttcttcttcttcttcttcttcttct 915
Db 125 CCCCACGTCACATTCCTACTCCGTTTCCATCCCTGGACCCAGGTTGGCCAACTCT 66
Oy 916 cctgttaagaaccagagaaactatatttagctctgtgtgcatatggtctcaatc 975
Db 65 CCTGTAAGAACACAGAGAACTATTGAGGCTGTGTGTCATATGCTCAGTCACA 6
```

```
RESULT 5
PCT-US01-15674A-421/C
: Sequence 421, Application PC/TUS0115674A
: GENERAL INFORMATION:
: APPLICANT: JOAN, KNOUL H
: APPLICANT: ROGAN, PETER K
: TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT
: FILE REFERENCE: 30307
: CURRENT APPLICATION NUMBER: PCT/US01/15674A
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 613
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 421
: LENGTH: 338
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: repeat_region
: LOCATION: (1)..(338)
: OTHER INFORMATION: cheshire_b
PCT-US01-15674A-421
```

Query Match 4.1%; Score 93; DB 1; Length 338;  
Best Local Similarity 65.8%; Pred. No. 4.7e-15;  
Matches 177; Conservative 1; Mismatches 66; Indels 25; Gaps 2;

```
Oy 900 gttggggaactcttcttgaagaaccagagaaactattttaggctctgtgtcga 959
Db 335 GGTGTCGCAAACTTTTGTGTAAGGCGCAGATGTAATATTTTAGGCTTGGCGGCA 276
Oy 960 tatgtctcagtcacaactactctctctctctctctgttagcagaagaacttagcaaa 1019
Db 275 TAGGTCCTGTGCGCAACTACTCAACTCTGCGTTGTAAGCCAAAGACGACATAGACA 216
Oy 1020 tatgtcaacaacatattgtgacccat--gaaacttattattatagatagaaac 1077
Db 215 TATATAAAGATGGCGTGGCTGTGTTCCATTAACCTTATTTATGACACTGAAT 156
Oy 1078 tga-----aaatagtcttcttcttcttcttcttcttcttcttcttcttctt 1114
Db 155 TCAATTTTCATATATATTTTTCACATGTCACAAATATATCTTCTTTGATTTTCTTCCAAAC 96
Oy 1115 attaaaaacgtlaaaactactcttaggt 1143
Db 95 ATTAAATGTAAAAACCAATCTTAGT 67
```

```
RESULT 6
US-09-573-080A-421/C
: Sequence 421, Application US/09573080A
: GENERAL INFORMATION:
: APPLICANT: JOAN, KNOUL
: APPLICANT: ROGAN, PETER
: TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENER
: FILE REFERENCE: 30307
: CURRENT APPLICATION NUMBER: US/09/573,080A
: CURRENT FILING DATE: 2000-05-16
: NUMBER OF SEQ ID NOS: 479
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 421
: LENGTH: 338
: TYPE: DNA
: ORGANISM: Homo sapiens
```



PCT-US01-15674A-420

Query Match 3.8%; Score 86; DB 1; Length 224;

Best Local Similarity 75.4%; Pred. No. 3.7e-13;

Matches 107; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
QY 900 ggttggcaaacctctcgttaagaagacagagaactattttagctgtgtgcca 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 221 gggtcgccaaactttttctgtaaggccagatagtaatttttagcgttgccga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 960 tatgttcacagtcacacactatctctgtcctctgtgacgaagaactagcaaca 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 TAGGCTCTGTGCGCAACTCACTGCGCTTACGCCCAAGACCATATAGACAA 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1020 tatgtcacacaacatattgtgac 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 101 TATATAACGAATGGCGTGGC 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 10

US-09-573-080A-420/C

Sequence 420, Application US/09573080A

GENERAL INFORMATION:

APPLICANT: JOAN, KNOEL

APPLICANT: ROGAN, PETER

TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI

FILE REFERENCE: 30307

CURRENT APPLICATION NUMBER: US/09/573,080A

CURRENT FILING DATE: 2000-05-16

NUMBER OF SEQ ID NOS: 479

SOFTWARE: Patent version 3.0

SEQ ID NO 420

LENGTH: 224

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: repeat\_region

LOCATION: (1)..(224)

OTHER INFORMATION: cheshire\_e

PUBLICATION INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: Jurka, J; Walichewicz, J; Milosavljevic, A

TITLE: Prototypic sequences for human repetitive DNA

JOURNAL: Journal of Molecular Evolution

VOLUME: 35

ISSUE: 4

PAGES: 286-291

DATE: 1992-10-

DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)

DATABASE ENTRY DATE: 1996-01-26

DATABASE ENTRY DATE: 1996-01-26

US-09-573-080A-420

Query Match 3.8%; Score 86; DB 22; Length 224;

Best Local Similarity 75.4%; Pred. No. 3.7e-13;

Matches 107; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
QY 900 ggttggcaaacctctcgttaagaagacagagaactattttagctgtgtgcca 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 221 gggtcgccaaactttttctgtaaggccagatagtaatttttagcgttgccga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 960 tatgttcacagtcacacactatctctgtcctctgtgacgaagaactagcaaca 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 TAGGCTCTGTGCGCAACTCACTGCGCTTACGCCCAAGACCATATAGACAA 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1020 tatgtcacacaacatattgtgac 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 101 TATATAACGAATGGCGTGGC 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

US-60-243-468-358

Sequence 358, Application US/60243468

GENERAL INFORMATION:

APPLICANT: Beasley, Ellen

TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,

NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,

TITLE OF INVENTION: AND USES THEREOF

FILE REFERENCE: C0000929

CURRENT APPLICATION NUMBER: US/60/243,468

CURRENT FILING DATE: 2000-10-27

NUMBER OF SEQ ID NOS: 2121

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 358

LENGTH: 76534

TYPE: DNA

ORGANISM: HUMAN

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)..(76534)

OTHER INFORMATION: n = A,T,C or G

US-60-243-468-358

Query Match 3.8%; Score 86; DB 63; Length 76534;

Best Local Similarity 65.8%; Pred. No. 1.1e-11;

Matches 171; Conservative 0; Mismatches 85; Indels 4; Gaps 3;

```
QY 888 tgcctgacccagtggtggcaaacctctctgttaagaagacagagaactatttag 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29012 tgtgaagacaggtctgtgcaaacctttctgttaagaagacagagaactatttag 29071
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 948 ctctgtgccaatagtcctgaacacacacacacacacacacacacacacacacac 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29072 ctcaagggcactctgtgtcgcacacacacacacacacacacacacacacacac 29130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1008 aattgacaacatattgcaacaacatattg--accacatgaacattattatc 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29131 accctgacagacgatatataacacagtgatgtgtgttcaataaattttatg 29190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1066 gatacggaaacctgaataatgt-cttctcttgatttttccccaatcaataaaac 1124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29191 aatactgaatttgatttatataaattatttgatttttccccaacttaaaa 29250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1125 gtaaaactactcttagtc 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29251 tgcaaaactattcttagtc 29270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

US-60-226-176-1401

Sequence 1401, Application US/60226176

GENERAL INFORMATION:

APPLICANT: Ring, Huijun Z.

APPLICANT: Malsen, Gareth

APPLICANT: Townley, David

TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes

FILE REFERENCE: CX-0013-1 P

CURRENT APPLICATION NUMBER: US/60/226,176

CURRENT FILING DATE: 2000-08-16

NUMBER OF SEQ ID NOS: 2447

SOFTWARE: PERL Program

SEQ ID NO 1401

LENGTH: 179040

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: GB:AC022732\_3

US-60-226-176-1401

Query Match

3.6%; Score 83.2; DB 61; Length 179040;

Best Local Similarity -63.28; Pred. No. 1.1e-10;  
Matches 144; Conservative 0; Mismatches 83; Indels 1; Gaps 1

[illegible]

RESULT 13  
NE-60-223-

```

1  ? Sequence 1401, Application US/602333468
2  ? US-60-233-468-1401
3  ?
4  ? GENERAL INFORMATION:
5  ?
6  ? APPLICANT: Ring, Huijun Z.
7  ?
8  ? APPLICANT: Malsen, Gareth
9  ?
10 ? APPLICANT: Towmley, David
11 ?
12 ? APPLICANT: Morris, Macdonald
13 ?
14 ? APPLICANT: Valdes, Ana
15 ?
16 ? TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated With ADME Genes
17 ?
18 ? FILE REFERENCE: GX-001332 P
19 ?
20 ? CURRENT APPLICATION NUMBER: US/60/233,468
21 ?
22 ? CURRENT FILING DATE: 2000-09-18
23 ?
24 ? NUMBER OF SEQ ID NOS: 2488
25 ?
26 ? SOFTWARE: PERL Program
27 ?
28 ? SEQ ID NO 1401
29 ?
30 ? LENGTH: 179040
31 ?
32 ? TYPE: DNA
33 ?
34 ? ORGANISM: Homo sapiens
35 ?
36 ? FEATURE:
37 ?
38 ? NAME/KEY: misc_feature
39 ?
40 ? OTHER INFORMATION: GB:AC022732_3
41 ?
42 ? US-60-233-468-1401

```

Query Match	3.68;	Score 83.2;	DB 62;	Length 179040;
Best Local Similarity	63.28;	Pred. No. 1.1e-10;		
Matches 144;	Conservative	0;	Mismatches 83;	Indels 1;
				Gaps 1

[illegible]

RESULT 14  
US-60-313-

```

; sequence 1401, Application US/6031337T
; GENERAL INFORMATION:
; APPLICANT: Ring, Huijun Z.
; APPLICANT: Malsen, Gareth
; APPLICANT: Townley, David

```

```

1  APPLICANT: Moritt, MacDonald
2  TITLE OF INVENTION: Single Nucleotide Polymorphisms Associated with ADME Genes
3  FILE REFERENCE: GX-0013-5 P
4  CURRENT APPLICATION NUMBER: US/60/313,371
5  CURRENT FILING DATE: 2001-08-16
6  NUMBER OF SEQ. ID NOS.: 2447
7  NUMBER OF

```

Query Match	3.68;	Score 83.2;	DB 70;	Length 179040;
Best Local Similarity	63.28;	Pred. No. 1.1e-10;		
Matches 144;	Conservative 0;	Mismatches 83;	Indels 1;	Gaps 1;

Qy	879	cglttcccaagcctcgtgaccccaaggttttggcaaacctcttcctctgtaagaagcaacgcagac	938
Db	5023	cttggagaagccttaggaacgggggtcaacaacaccttttttgtaagaacacgaagttaa	5082
Qy	939	tattttaggcctctgtctgtgcacatacgtctcagtcacaaactactcatctctgcctctgtag	998
Db	5083	tatttcaggttcttgcaagccacatacgtctctctcgtctgtaacatgctgcacacccctgcaggtag	5142
Qy	999	cacgaaagcgaattagcaacacatatgtcacaacaaacatatgtgagacccatgaaactttt	1054
Db	5143	caccaaagccagtcatacagacata-ctcaagcgaaagagagctgtcctaataaaattttatc	5201
Qy	1059	tattatgatacagaagaaacccgaaataatagctctctttctttgatttttt	1106
Db	5202	tatgactcgaaatttgaaatttcataacttttagtttctttcttttttt	5249

RESULT 15  
US-60-212-

```

: Sequence 256, Application US/60212664
: GENERAL INFORMATION:
: APPLICANT: ladunga, Steve
: APPLICANT: Spier, Gene
: APPLICANT: Greenberg, Simon
: APPLICANT: Radkai, Steven
: APPLICANT: Wang, Yu
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
: TITLE OF INVENTION: USBS THEREOF
: FILE REFERENCE: CLO00067
: CURRENT APPLICATION NUMBER: US/60/212,664
: CURRENT FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 636
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 256
: LENGTH: 163174
: TYPE: DNA
: ORGANISM: HUMAN
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(163174)
: OTHER INFORMATION: n = A,T,C or G
US-60-212-664-256

```

Query Match	3.68;	Score 83;	DB 60;	Length 163174;
Best Local Similarity	65.28;	Pred. No. 1.2e-10;		
Matches 122;	Conservative	0;	Mismatches 65;	Indels 0;
			Gaps	0;

QY 903 tggcgcaactcttccctgtaagaagcagacaggaactatlttaagctcgtgtgcataat 962  
| | | | | | | | | | | | | | | | | | | | | |  
DB 92341 TCGGCAACTTTCCTGTAAAGGCTCAGATAGTAATAATTATTTAGGCTTTGTGGGCACAG 922B2





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 05:14:32 ; Search time 97.23 Seconds  
(without alignments)  
5658.943 Million cell updates/sec

Title: US-09-515-369B-1\_COPY\_1\_2240  
Perfect score: 2240  
Sequence: 1 taatacgaactcaactataggg.....aaatccacaagaagccttgc 2240

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents, NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCtus.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71.8	3.2	7218	1 US-08-232-463-14	Sequence 14, Appl
2	67.6	3.0	6235	4 US-09-305-384-5	Sequence 5, Appl
3	67.6	3.0	6679	4 US-09-305-384-1	Sequence 1, Appl
4	62.2	2.8	178	4 US-08-991-789A-210	Sequence 210, App
5	62.2	2.8	178	4 US-09-063-451-210	Sequence 14, Appl
6	52.6	2.3	7218	1 US-08-232-463-14	Sequence 14, Appl
7	52	2.3	3786	4 US-08-975-762-42	Sequence 42, Appl
8	52	2.3	3786	4 US-09-295-028-42	Sequence 42, Appl
9	52	2.3	3786	4 US-09-106-582-42	Sequence 42, Appl
10	50.6	2.3	502	2 US-08-967-101-91	Sequence 91, Appl
11	50.6	2.3	502	2 US-08-592-541-91	Sequence 91, Appl
12	50.6	2.3	502	3 US-09-124-698-91	Sequence 91, Appl
13	50.6	2.3	502	4 US-09-127-480-91	Sequence 91, Appl
14	50.6	2.3	502	4 US-08-496-841C-91	Sequence 91, Appl
15	50.6	2.3	1883	2 US-08-967-101-10	Sequence 10, Appl
16	50.6	2.3	1883	2 US-08-967-101-153	Sequence 153, Appl
17	50.6	2.3	1883	2 US-08-592-541-10	Sequence 10, Appl
18	50.6	2.3	1883	2 US-08-592-541-153	Sequence 153, Appl
19	50.6	2.3	1883	3 US-08-888-077A-8	Sequence 8, Appl
20	50.6	2.3	1883	3 US-09-124-698-10	Sequence 10, Appl
21	50.6	2.3	1883	3 US-09-124-698-153	Sequence 153, Appl
22	50.6	2.3	1883	4 US-09-127-480-10	Sequence 10, Appl
23	50.6	2.3	1883	4 US-09-127-480-153	Sequence 153, Appl
24	50.6	2.3	1883	4 US-08-496-841C-10	Sequence 10, Appl
25	50.6	2.3	1883	4 US-08-496-841C-153	Sequence 153, Appl
26	48.2	2.2	80246	4 US-09-078-294-4	Sequence 4, Appl
27	48.2	2.2	80595	4 US-09-078-294-3	Sequence 3, Appl

## ALIGNMENTS

28	45.8	2.0	1734	2 US-08-858-052-2	Sequence 2, Appl
29	45.8	1.9	1734	2 US-08-200-284-2	Sequence 2, Appl
30	43.6	1.9	460	2 US-08-967-101-48	Sequence 48, Appl
31	43.6	1.9	460	2 US-08-592-541-48	Sequence 48, Appl
32	43.6	1.9	460	3 US-08-124-698-48	Sequence 48, Appl
33	43.6	1.9	460	4 US-09-127-480-48	Sequence 48, Appl
34	43.6	1.9	460	4 US-08-496-841C-48	Sequence 48, Appl
35	42.8	1.9	59065	4 US-09-813-817-3	Sequence 3, Appl
36	42.4	1.9	289	4 US-09-007-005-17	Sequence 17, Appl
37	42.4	1.9	289	4 US-09-244-796-17	Sequence 17, Appl
38	41.4	1.8	703	4 US-09-313-300-6	Sequence 6, Appl
39	41.4	1.8	176373	3 US-09-128-155-17	Sequence 17, Appl
40	41.2	1.8	7505	4 US-09-078-294-13	Sequence 13, Appl
41	39.8	1.8	350	2 US-08-332-766A-20	Sequence 20, Appl
42	39.2	1.8	84495	4 US-09-797-906-3	Sequence 3, Appl
43	39.2	1.8	152331	3 US-09-128-155-16	Sequence 16, Appl
44	38.6	1.7	43795	3 US-08-742-185-101	Sequence 101, App
45	38.2	1.7	1732	1 US-08-480-346-2	Sequence 2, Appl

RESULT 1  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFLINER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOMPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)863-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pl-F1s  
US-08-232-463-14





Oy 544 ctaaacacacatggtggaaggggggagtgaggagagtgctgtgtgccc 603  
Db 1430 RR 1371  
Oy 604 taacttggagtgtaagactcggccctccaggaacaaagattcagctgtgtgagct 663  
Db 1370 RR 1311  
Oy 664 atagccagagacgtcgtgcagagagtgcaagagatatttctgtcttaagaaat 723  
Db 1310 RR 1251  
Oy 724 aaacacacactgagtgagagaggggtgtgtgtccagagagatgtggagag 783  
Db 1250 RR 1191  
Oy 784 tctgcacaggg 794  
Db 1190 RRRRRRRRRRR 1180

RESULT 7  
US-08-975-762-42  
Sequence 42, Application US/08975762  
Patent No. 6207169  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,762  
FILING DATE: 21-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-622-4500  
TELEFAX: 206-682-6031  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3786 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-975-762-42

Query Match 2.3%; Score 52; DB 4; Length 3786;  
Best Local Similarity 67.2%; Pred. No. 8.7e-06;  
Matches 90; Conservative 0; Mismatches 40; Indels 4; Gaps 1;

Oy 914 ttccgtgaagaaacagagaaactatttagctctgtgtgcacatagtc---tca 969  
Db 3296 ttctgtgtacgtccagacagtaaacattttggactttgtatgtatgtgtcttct 3355  
Oy 970 gtcaaacactactatcttgcctctgttagcagaaagaactatgaacaataatgtcaaca 1029

Db 3356 GTTCAACTATGTAACCTCTTCANTGTAGACAGCAAGCGCGTGCAGACAAATATGTAAACA 3415  
Oy 1030 aacataatgtgacc 1043  
Db 3416 GATGAGCATGACTC 3429  
RESULT 8  
US-09-295-028-42  
Sequence 42, Application US/09295028  
Patent No. 6277381  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS  
AND TREATMENT OF EHRLICHIA INFECTION  
FILE REFERENCE: 210121.439c4  
CURRENT APPLICATION NUMBER: US/09/295,028  
CURRENT FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 42  
LENGTH: 3786  
TYPE: DNA  
ORGANISM: Ehrlichia sp.  
US-09-295-028-42

Query Match 2.3%; Score 52; DB 4; Length 3786;  
Best Local Similarity 67.2%; Pred. No. 8.7e-06;  
Matches 90; Conservative 0; Mismatches 40; Indels 4; Gaps 1;  
Oy 914 ttccgtgaagaaacagagaaactatttagctctgtgtgcacatagtc---tca 969  
Db 3296 ttctgtgtacgtccagacagtaaacatttggactttgtatgtatgtgtcttct 3355  
Oy 970 gtcaaacactactatcttgcctctgttagcagaaagaactatgaacaataatgtcaaca 1029  
Db 3356 gtgcgaactactgacctcttcattctgttagcagaaagcgcgacacataatgtcaaca 3415  
Oy 1030 aacataatgtgacc 1043  
Db 3416 gatgagcatgactc 3429

RESULT 9  
US-09-106-582-42  
Sequence 42, Application US/09106582  
Patent No. 6306402  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,582  
FILING DATE: 29-JUN-1998

THERAP

```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 3786 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-106-582-42

```

```

Query Match      2.3%; Score 52; DB 4; Length 3786;
Best Local Similarity 67.2%; Pred. No. 8.7e-06;
Matches 90; Conservative 0; Mismatches 40; Indels 4; Gaps 1;

```

```

Qy 914 ttccgttaaaagacacagagacattttagctctgtgtgcacatagtc---tca 969
Db 3296 TTTGTGTGTACTGACAGAGTAACATTGTTGACTTTGTATGTATGTGCTCTTCT 3355
Qy 970 gtcaaacatctatctctctctgtgacagaaagcaatagcaacataatgtcaaa 1029
Db 3356 GTTGCACACTACTGAACTCTTCATTTGTACAGACAGCGCTGACAGCAATATGTAAACA 3415
Qy 1030 aacatatgtgacc 1043
Db 3416 GATGAGCATGACCTC 3429

```

```

RESULT 10
US-08-967-101-91
Sequence 91, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 502 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-967-101-91

```

```

Query Match      2.3%; Score 50.6; DB 2; Length 502;
Best Local Similarity 64.3%; Pred. No. 6.8e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

```

```

Qy 917 ctgttaaagacacagacagacattttagctctgtgtgcacatagtcctcagtcaca 976
Db 264 CTGTAAAGAGCCAGACGCTG--AATATTTTAAAGCTATGCTATGCTCCAGGGCAA 321
Qy 977 ctactcatctctctctctgtgacagaaagcaatagcaacataatgtcacaacacat 1036
Db 322 ACACCTCAACTGCTGCTACTGTAGTGTGAAGCAGCAGACACATGTATTAAACAGGAG 381
Qy 1037 gtgaccacatgaaaacttatt 1059
Db 382 GTGTCACCTTCCATGAAAGTT 404

```

```

RESULT 11
US-08-592-541-91
Sequence 91, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-592-541-91

```

```

Query Match      2.3%; Score 50.6; DB 2; Length 502;
Best Local Similarity 64.3%; Pred. No. 6.8e-06;
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

```

```

Qy 917 ctgttaaagacacagacagacattttagctctgtgtgcacatagtcctcagtcaca 976
Db 264 CTGTAAAGAGCCAGACGCTG--AATATTTTAAAGCTATGCTATGCTCCAGGGCAA 321

```

OY	977	cctccatctcgtccctctgagccgaagaacatttagcaaacatgatgctcacaacatat	1036
Dd	322	AACCTCAACTGTGGTACTGCTAGTGSTGAAGAAGCAGCACAGAAATCTTTAACCCAGGAG	381
OY	1037	gtgaccoccatgaaaccttatttc	1059
Dd	382	GTCGTACTTTCATCAAGATT	404

RESULT 12  
US-09-124-

```

US-09-124-698-91
: Sequence 91, Application US/09124698
: Patient No. 611978
:
: GENERAL INFORMATION:
: APPLICANT: ST. GEORGE-HYSLOP, PETER H
: APPLICANT: ROMMENS, JOHANNA M
: APPLICANT: FRASER, PAUL E
: TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
: TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
: NUMBER OF SEQUENCES: 183
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TESTA, HUMMELT & THIBEAULT
: STREET: High Street Tower - 125 High Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/124,698
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/592,541
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Pletcher, Edmund R.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 91:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 502 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
:
US-09-124-698-91

```

Query Match	2.38;	Score 50.6;	DB 3;	Length 502;
Best Local Similarity	64.3%;	Pred. No. 6.8e-06;		
Matches	92;	Conservative	0;	Mismatches 49;
				Indels 2;
				Gaps 1

## RESULT 13

US-09-127-480-91  
; Sequence 91, Application US/09127480  
; Patent No. 6104153

Query Match	2.38;	Score 50.6;	DB 4;	Length 502;
Best Local Similarity	64.38;	Pred. No. 6.8e-06;		
Matches 92;	Conservative	0;	Mismatches 49;	Indels 2;
				Gaps 1;

RESULT 14  
 US-08-496-841C-91  
 ; Sequence 91, Application US/08496841C  
 ; Patent No. 6210919  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ST. GEORGE-HYSLOP, PETER H  
 ; ROMMENS, JOHANNA M  
 ; FRASER, PAUL E  
 ; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
 ; TO ALZHEIMER'S DISEASE  
 ; NUMBER OF SEQUENCES: 175  
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby, PC  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,841C  
FILING DATE: 28-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul F. Fehner, Ph.D.  
REGISTRATION NUMBER: 35,135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237

INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 91:

US-08-496-841C-91

Query Match 2.3%; Score 50.6; DB 4; Length 502;  
Best Local Similarity 64.3%; Pred. No. 6.8e-05;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgttaagaacagacgaactatttaggtctgtgtccatattgtcagtcacaa 976  
DB 264 CTGTAAAGACGACGCTG--AATATTTTAACTCTATGCTCTCCAGGGCAA 321  
QY 977 ctactacatctgtcctctgtgacgaagaatlagcaacatatgtcacaacatat 1036  
DB 322 AACTCAACTGCTGCTAGTGTGAAGGACGACGACATGTTATTAACCAAGGAG 381  
QY 1037 gtgacccatgaactattt 1059  
DB 382 GTGCTCACTTCCATGAAGTT 404

RESULT 15  
US-08-967-101-10/C  
Sequence 10, Application US/08967101  
Patent No. 5840540  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURNITZ & THIBEAULT  
STREET: High Street tower - 125 High Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/967,101  
FILING DATE: 10-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/592,541  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plicher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1883 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-967-101-10

Query Match 2.3%; Score 50.6; DB 2; Length 1883;  
Best Local Similarity 64.3%; Pred. No. 1.5e-05;  
Matches 92; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 917 ctgttaagaacagacgaactatttaggtctgtgtccatattgtcagtcacaa 976  
DB 239 CTGTAAAGACGACGCTG--AATATTTTAACTCTATGCTCTCCAGGGCAA 182  
QY 977 ctactacatctgtcctctgtgacgaagaatlagcaacatatgtcacaacatat 1036  
DB 181 AACTCAACTGCTGCTAGTGTGAAGGACGACGACATGTTATTAACCAAGGAG 122  
QY 1037 gtgacccatgaactattt 1059  
DB 121 GTGCTCACTTCCATGAAGTT 99

Search completed: August 16, 2002, 05:15:29  
Job time: 15898 sec

Mon Aug 19 06:46:40 2002

us-09-515-369b-1\_copy\_1\_2240.rn1